TOWARDS A SMARTER LIGHT SHEET MICROSCOPE

JIAYE HE
geboren am 4. Januar 1992, in Harbin (China)

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium (dr.rer.nat.)

der Fakultät Mathematik und Naturwissenschaften
der Technischen Universität Dresden

gutachter: Prof. Stephan Grill, Technische Universität Dresden
Prof. Kevin Eliceiri , University of Wisconsin, Madison
betreuer: Dr. Jan Huisken, Morgridge Institute for Research
eingereicht am: 15. Juli 2019
verteidigt am: 2. Oktober 2019

E-mail: jhe@morgridge.org

ABSTRACT

Selective plane illumination microscopy (SPIM) is becoming the method of choice for long-term 3D fluorescence imaging thanks to its low photo-toxicity and high imaging speed. However, SPIM is very data intensive: A single SPIM experiment can easily generate terabytes of image data, which is often overwhelming for biologists to handle. Moreover, large SPIM datasets often require additional computational power for processing. There is a lack of optimized analysis software to visualize and quantify such large datasets. As a result, the data size burden is limiting the accessibility of this immensely powerful technology.

In this thesis, I investigated the root of the data burden in SPIM. I found that although the raw data volume generated by SPIM is large, the data product after processing is often very small in comparison. As a result, there are two ways of alleviating the data burden: the raw data to data product conversion ratio can be improved and the process of converting raw data to data product can be streamlined. In my Ph.D. project, I demonstrated three different approaches to tackle the data burden in SPIM.

Firstly, I tested different lossless data compression methods on standard SPIM datasets that I collected during my thesis work. I found that integer compression algorithms are ideal for SPIM images. The data size can be reduced by more than half when compressed losslessly on-the-fly, reducing storage stress.

Secondly, if the image quality could be improved, raw images would contain a higher amount of useful information and it would be less wasteful to store the large dataset. To illustrate this concept, I created a custom on-the-fly image analysis software that automatically selects the optimal imaging view in a multi-view SPIM experiment. By applying the workflow to zebrafish embryo imaging, I showcase that each multi-view dataset contains more information than in the conventional case. Moreover, it became possible to reduce the number of imaging views without compromising data quality.

Lastly, raw data can also be converted into data product on-the-fly. The need to store raw images is often a result of the disconnect between imaging and data analysis. If the raw data can be analyzed in memory as soon as they are captured by the microscope, there is no need to keep the raw image data. I have built a custom image analysis pipeline to quantify zebrafish Rohon-Beard cells' axon branching patterns. The image analysis software semi-automatically performs sample surface extraction and image unwrapping. The resultant dataset is a flat lateral view of the embryo. The processed dataset is less than 10% the size of the original images. I also show that through directionality analysis, the processed data can be used to identify wild type embryos from drug-treated samples.

I also showcase a couple of other custom SPIM imaging workflows that I helped create. I have imaged patient-derived cancer spheroids and *Xenopus* oocytes in collaboration with other researchers. Here, the smart microscopy concept helped facilitate many data processing challenges involved.

Overall, my thesis showcased that the data burden in SPIM can be addressed effectively by integrating image processing closely into the image capture process. I call this overall concept "smart microscopy" and I believe it is the future of fluorescence microscopy.

PUBLICATIONS

UNDER REVIEW

He, J. and Huisken, J. *Image quality guided smart rotation improves coverage in microscopy.* Nature Communications.

Campbell, C., Lancman, J.J., Palazon, R.S., Matalonga, J., He, J., Graves, A., Zeng, X.I., Huisken, J., Traver, D. and Dong, P. D. S. *In vivo lineage conversion of vertebrate muscle into early endoderm-like cells.* Science

Rosowski, E. E., He, J., Huisken, J., Keller, N. P. and Huttenlocher, A. *Efficacy of voriconazole against A. fumigatus infection depends on host immune function.* Antimicrobial Agents and Chemotherapy.

MANUSCRIPTS IN PREPARATION:

Favreau, P., He, J., Gil, D., Deming, D., Huisken, J. and Skala, M. *Label-free redox imaging of patient-derived spheroids using Selective Plane Illumination Microscopy.*

Haynes, E. M., He, J., Huisken, J. and Halloran, M. *Integrated mid-throughput screening platform for zebrafish axon branching defects.*

ACKNOWLEDGEMENTS

As I am sitting here writing the acknowledgment section of my Ph.D. thesis, I cannot believe four years have passed since I became a part of the Huisken lab. Looking back, this has been the most eventful few years of my life by far. I have met and worked with amazing scientists, traveled to places that I never expect to be visiting, and most importantly, learned how to be better as a person and as a researcher. I want to express my sincerest gratitude to the people that supported me over the past few years.

First of all, I would like to thank my Ph.D. mentor, *Jan Huisken*. I am always amazed by Jan's vision and creativity. The countless number of progress report meetings turned into brainstorming sessions. I always leave his office with more ideas than I enter with. At the same time, Jan allows me to be independent and pursue new ideas. The only thing I never agree with Jan is his love for Fuji cameras. Sony cameras are clearly better.

Secondly, I want to thank the members of the Huisken lab.

The first person in the Dresden Huisken lab that I grew close to is *Nico Scherf*. I still remember the tiny back room where Nico and I used to sit and program away. I am always impressed by the beautiful visualization he can generate and the amount of Club-Mate he can consume in an afternoon. I can always count on Nico to point me in the right direction when I am struggling with image analysis. I would also like to thank *Stephan Daetwyler*. We are usually the only people working late at the institute. Stephan always cheers me up with his encouraging words and positive attitudes during those late night programming sessions.

I am also very thankful to other members of the Dresden Huisken lab. *Wiebke Jahr* patiently taught me how to get started with optics. *Benjamin Schmid* taught me how to get started with LabView (which I still hate). I will always remember those board games nights at Triangle with all the lab members.

I would also like to thank *Nicola Gritti*. Nicola has been more than a colleague for me. We can talk about anything, from computer science to mischievous adventures. My life in Madison over the past years would not be the same without Nicola. The same can be said about *Anjalie Schlaeppi*. As we are the only Ph.D. students that moved to the US, we shared a lot of the struggles together, which made it easier for me to overcome the cultural and social barriers.

I would also like to thank *Kurt Weiss*. When I first moved to Madison, Kurt helped me settle in (during a snowstorm no less). Kurt showed me what mid-western hospitality actually means.

I would like to also thank all current and former Huisken lab members: *Alyssa, Anjalie, Heather, Joe, Kurt, Rory, Todd, Benjamin, Frederic, Gloria, Gopi, Isabell, Max, Michael, Michaela, Nico, Nicola, Stephan, and Wiebke*. You created such a wonderful and enjoyable working environment. I will always cherish my memory in the lab.

I cannot thank the Morgridge IT team enough. Without a reliable IT infrastructure, there is no smart microscopy. *Neil* and *Ben* set up the perfect IT pipeline for the Madison Huisken lab and me. I can always rest easy knowing my data is in good hands. I would also like to thank *Alyssa*, who helped me with so many fish husbandry tasks.

I would like to thank all of my amazing collaborators in Madison: *Ani, Emily, Veronika, Peter, Richard,* and so many more. I am always amazed by their research, and I am so proud to be able to contribute.

When I first started imaging with *Liz*, I was quickly dazzled by her enthusiasm in microscopy and biology. Our collaboration goes beyond me simply providing the instrument I constructed. Each imaging session was filled with exciting exchanges about biology and technology. I am proud to be able to call Liz my friend, as she is one of the most inspiring scientists I know.

Of course, I have to thank all my friends: *Yong* for always being there for me even though we don't see each other often enough; *Tzi-yang* for always able to put a smile (or facepalm) on my face and *Niklas* for all the misbehaving we do.

I would like to also thank my thesis advisory committee: *Ivo Sbalzarini* and *Suzanne Eaton*. Ivo has provided much-needed insights when I am struggling to navigate through the complex Bio-image analysis ecosystem. Suzanne has always been there to remind me to not only think about the tools I build but also if they can lead to biological discoveries. As I finish writing this thesis, I learnt that Suzanne has unfortunately passed away during a conference in Greece. My heart goes out to her family and friends.

Finally, I want to express my deepest gratitude and apology to my parents, *Hong Shi* and *Dongbo He*. It has been 11 years since I left home and unable to spend enough time with my family. Even though we only meet once a year (or less), without your constant encouragement and support, it would be impossible for me to carry on. I promise I will be home more often, I promise.

CONTENTS

3 SMART ROTATION WORKFLOW TO IMPROVE LIGHT SHEET SAMPLE COVERAGE 33

IV APPENDIX 97

List of Figures

List of Tables

Part I

INTRODUCTION

1

A smarter microscope

"Progress in science depends on new techniques, new discoveries, and new ideas, probably in that order."

— Sydney Brenner

1.1 Fluorescence microscopy as a discovery tool

Scientific discoveries often start with observations. Many scientists trace their earliest scientific inspiration to a specific common observation. Since the invention of the simple yet powerful compound microscope from the 17th century, scientists have been making better tools to observe the unobservable micro-cosmos (**Fig** 1.1).

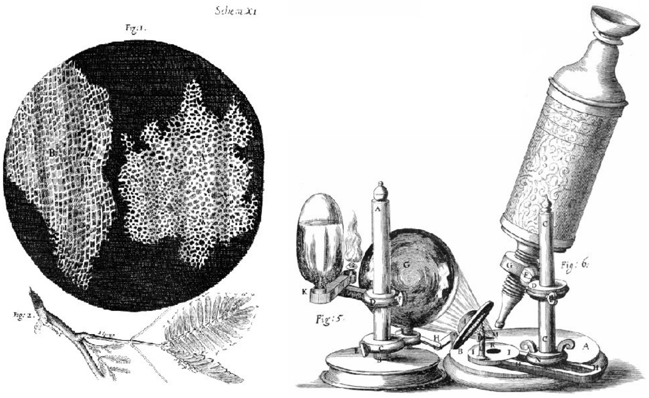


Figure 1.1 The first microscope circa the 1600s. Image from Wikipedia.

It was not until the discovery of fluorescent proteins in the second half of the 20th century that transformed optical microscopy into one of the most utilized tools in biology today. Over the following decades, innovations in optical design as well as protein engineering have made the fluorescence microscope one of the most powerful techniques for biological discoveries. The ability to observe and quantify some of the smallest building blocks of life in their spatial and temporal context is immensely valuable. Fluorescence microscopy has the advantage of performing the measurement in a non-destructive way when compared to other quantification techniques in biological science such as genome sequencing and mass spectrometry.

For the past two decades, the microscopists have pushed the horizon of what is physically possible in fluorescence microscopy. Thanks to the development of optical sectioning techniques such as confocal microscopy and two-photon excitation fluorescence microscopy, the observation of an entire living sample in three dimensions is more accessible than ever. The invention of super-resolution techniques such as Stimulated Emission Depletion microscopy (STED, Hell and Wichmann, 1994) and Photoactivated Localization Microscopy (PALM, Betzig et al., 2006) has pushed the spatial resolution limit of optical microscopy beyond the diffraction limit, which was postulated to be an impossible feat. With the help of sample clearing (Erturk and Bradke, 2013) and expansion microscopy (ExM, F. Chen et al., 2015), the entire fly brain has been imaged with a sub-micron resolution at speeds thousands of time faster than electron microscopy (Gao et al., 2019). There have even been efforts to perform spatial RNA transcriptomics with only fluorescence imaging using custom fluorescence in situ hybridization (FISH, K. H. Chen et al., 2015).

1.2 STANDARD WORKFLOW IN BIO-IMAGING EXPERIMENTS

The evolution of fluorescence microscopy techniques comes with a rising requirement for the level of expertise in different aspects of the imaging experiment. To be able to utilize the power of advanced imaging methodologies fully, researchers need to pay specific attention to how to build a successful bio-imaging workflow.

1.2.1 BIOLOGIST DRIVEN WORKFLOW

Fluorescence images give a direct visual representation of the underlying process and can therefore already answer a wealth of questions. Fluorescence microscopy is also often used only as a phenotyping method to verify findings from other methods. A lot of bio-imaging experiments are designed to only confirm binary hypothesis such as the expression of a certain protein in a process. Therefore no sophisticated quantification is needed beyond a simple binary labeling or manual annotation. In this scenario, biologists take the helm in all parts of the workflow:

- **Biologists** come up with a hypothesis.
- **Biologists** perform imaging experiment.
- **Biologists** analyze the images and obtain the final results.

In my opinion, fluorescence microscopy is under-utilized as a quantitative method for biological discovery in this case. However, utilizing more sophisticated image analysis tools for more complex quantification is not an easy task. Unlike other bioinformatics tools such as genome sequencing and mass spectrometry, the data output of fluorescence images cannot be easily standardized. Each new experiment may impose completely different image analysis challenges. For example, if one wishes to track cells during zebrafish gastrulation, the specific labeling method can have profound impacts on the way images need to be analyzed. Nucleus marker visualizes a specific region within the cell, which is more beneficial for tasks such as cell tracking. On the other hand, membrane markers make cell segmentation easier and also offer insight into cell interactions. The two labeling methods require completely different segmentation and tracking methods.

Although there are already many open-source and commercial softwares available to biologists, their performance often depends on many factors specific to each imaging experiment. For example, the signal to noise ratio (SNR) of an image depends on the brightness of the fluorophores as well as the microscope. SNR has a significant impact on the performance of certain histogram based image analysis algorithms such as watershed segmentation and automatic thresholding. Existing tools usually need to adapt to specific tasks for optimal performance. A computer vision expert is often needed to facilitate more involved analysis tasks.

1.2.2 Bio-image analysis

The field of image-based bioinformatics aims to unlock the full potential of fluorescence microscopy and extract quantitative information from images. Thanks to the great open source community, tools such as *Fiji* and *CellProfiler* exist to provide a playground for bio-image analysts to try out different methods (Carpenter et al., 2006;

Schindelin et al., 2012). Users can contribute to this community effort by building their own tools into plugins for these platforms. However, the plugins varies greatly in terms of versatility and user-friendliness. In some more user-friendly plugins, the software is highly automatic and can perform well even without parameters fine-tuning. If users are interested in optimizing the performance of the plugin, these software offer extensive configurability and guide the user through the optimization process (Hörl et al., 2018). More frequently, image analysts have to modify the existing plugins or even create new ones to address a specific task. An image analyst assisted imaging workflow becomes:

- **Biologist** comes up with a hypothesis.
- **Biologist** performs imaging experiment.
- **Biologist** brings data to image analyst.
- **Image analyst** builds the analysis software to obtain quantitative results.
- **Biologist** interprets the final results.

It is worth noting that in this scenario, the image analysts' treatment of image data is post-mortem as they do not contribute to how data is generated. Therefore the sufficiency of the information contained in the image has to be guaranteed by the one experimenting (biologists). To alleviate such conundrums, biologists can talk to the image analyst beforehand and specify the experimental parameters, such as the spatial-temporal resolution, together.

1.2.3 Integrated workflow

One other critical role that was yet been left out of the equation is the microscopist. Conventionally, biological institutes house commercially available instruments and microscopists are tasked only to maintain the instrument and create a standard operating procedure for the end-users. However, scientific discoveries most often rely on cutting edge technologies. The commercialization of new ideas takes a long time. For example, the first commercially available light sheet system from Zeiss became available roughly eight years after the publication of the first Selective Plane Illumination Microscope (SPIM, Huisken et al., 2004). Many labs, therefore, home-built their light sheet microscopes to gain access to the technology early (**Fig** 1.2). In this case, the microscope developer has the full freedom to design a system custom tailored to different applications. Therefore they should also be included in the discussion of how

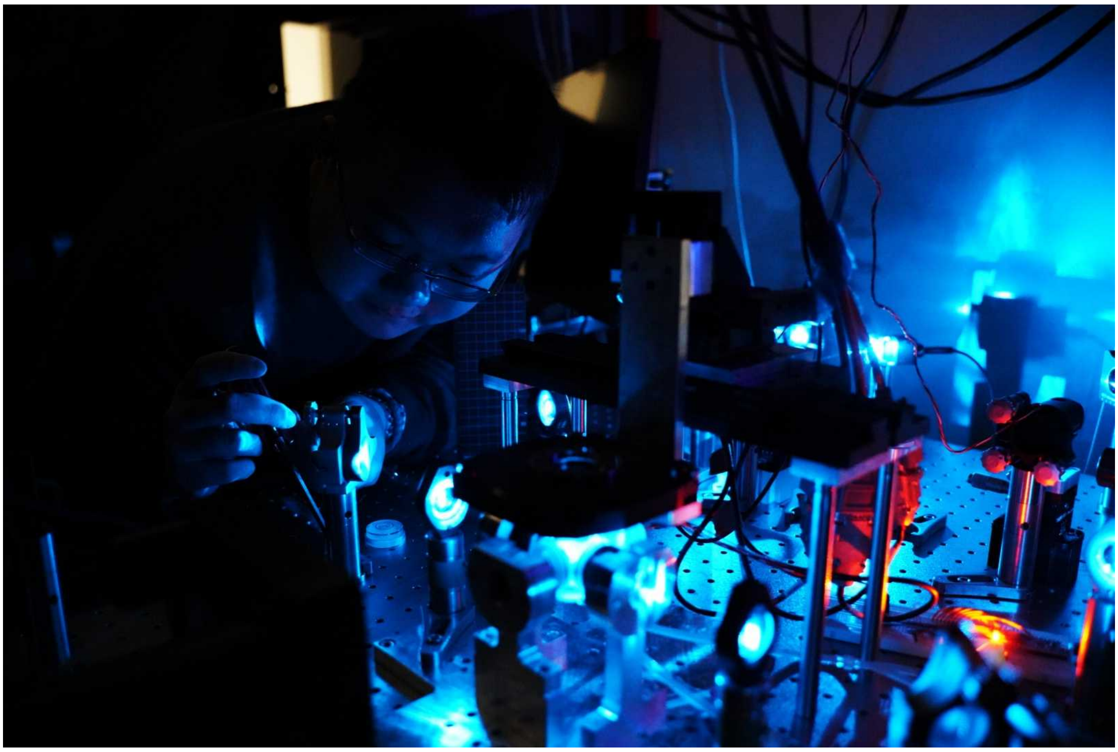


FIGURE 1.2 DIY microscope. Microscopists can custom make microscopes for different applications

the workflow should look like and what kind of images are required to make the data better for the quantification purposes.
The revised workflow for a bio-imaging experiment utilizing a home-built custom microscope is as follows.

- **Biologist** comes up with a hypothesis.
- **Biologist and Image analyst** discuss what kind of imaging is needed to answer the question.
- **Microscopist** builds a microscope capable of performing the experiment.
- **Microscopist/Biologist** performs the experiment.
- **Microscopist/Biologist** brings the data to the Image analyst.
- **Image analyst** quantifies the images to obtain final results.
- **Biologist** interprets the results.

Conventionally, the microscopist need to consider whether the spatial-temporal resolution of the microscope is sufficient for the specific process being observed. In addition, microscopists need to also consider if the imaging process would induce severe photo-damage that can affect the sample health. However, recent development in optical design and camera technology has brought the big data challenge to the microscopy field. Microscopists need to pay additional attention to how data is handled when utilizing new techniques such as light sheet fluorescence microscopy (LSFM). The imaging experiment bottleneck often is the data throughput rather than the limitation of optical instrumentation in the era of LSFM.

1.3 Light sheet fluorescence microscopy

For decades, the most popular three-dimensional fluorescence imaging technique has been confocal microscopy. In confocal microscopy, a tightly focused excitation beam is used to raster scan the sample to generate an image. Since the fluorophores above and below the plane are also illuminated, the fluorescence photons emitted by out-of-focus fluorophores are also collected by the objective lens. To eliminate the contribution from these out-of-plane signal in the final image, a pin-hole is placed in front of the detector. Since out-of-focus fluorescence light should not be focused at the detector's plane, they are rejected by the pin-hole. Only the signal from the plane of interest passes the pin-hole and is collected by the detector. As a result, confocal microscopy can achieve optical sectioning. However, since a significant amount of excitation light is wasted to illuminate out-of-focus fluorophores and only a fraction of the fluorescence signal is collected, confocal microscopy is inherently inefficient (**Fig** 1.4). Images generated by confocal microscopy usually have poor dynamic range and SNR. Moreover, during the generation of a single image stack, the same plane within the sample is illuminated multiple times, from both in-focus excitation light and out-of-focus exciation light. The sample therefore suffers from significant amount of photo-damage. Confocal experiments can usually run for a few hours as a result. In recent years, LSFM has gradually taken over as the method of choice for long-term three-dimensional imaging, thanks to its superior efficiency in excitation and photon collection.

The fundamental concept of using a light sheet to illuminate a thin volume within the sample to achieve optical sectioning can be traced back to Siedentopf and Zsigmondy, 1902. It was not until the last decade of the 20th century that the technique was used in bio-imaging applications. Orthogonal-plane fluorescence optical sectioning (OPFOS, Voie et al., 1993) were used to study the cochlea, and Thin Light Sheet Microscope (TLSM, Fuchs et al., 2002) were used to study bacteria. In 2004, *Huisken et*

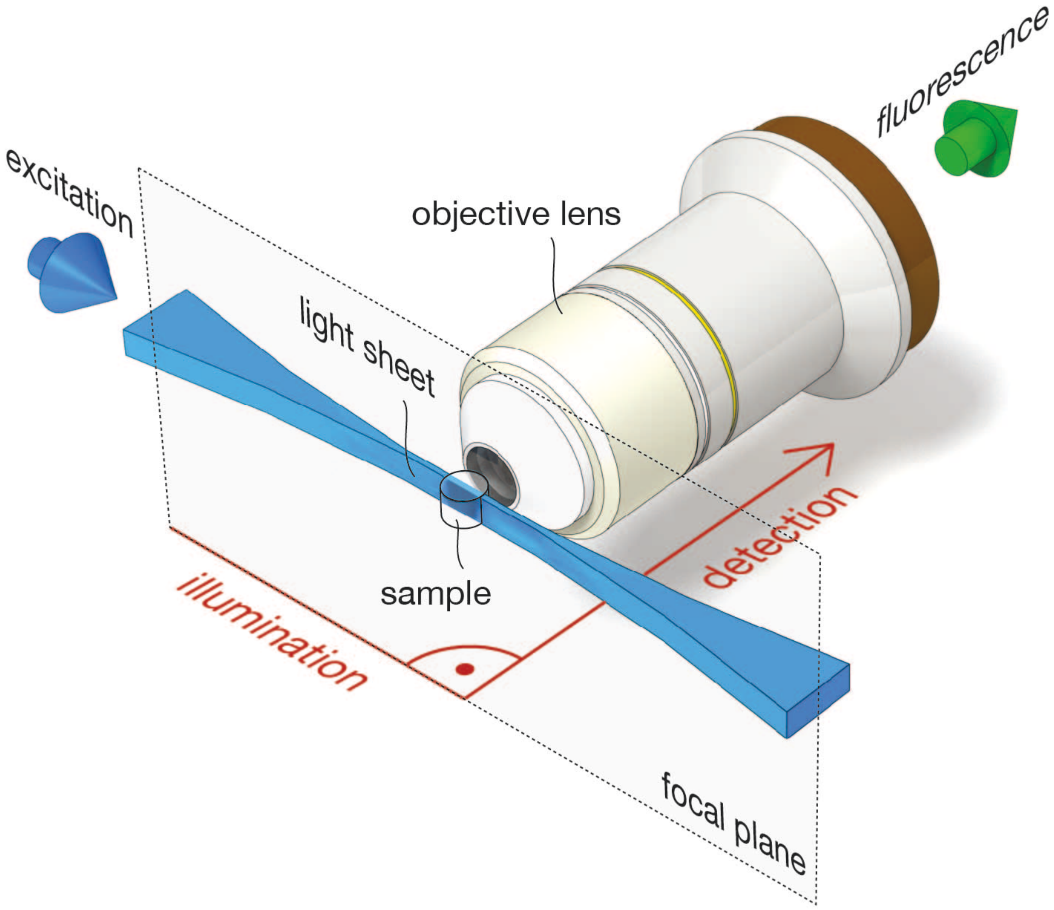


FIGURE 1.3 Principle of light sheet microscope. Adapted from Huisken and Stainier, 2009.

al. published the seminal paper that shaped how a modern light sheet microscope or Selective Plane Illumination Microscope (SPIM) is constructed. SPIM was first used to study the embryogenesis of Medaka fish and *Drosophila melanogaster* and later adapted to observe zebrafish development (Keller et al., 2008).

SPIM utilizes a cylindrical lens to shape the excitation light into a thin sheet. The sheet of light is projected through an objective lens to illuminate an entire plane within the sample at once. The plane of illumination coincides with the focal plane of a second objective that is used to collect the fluorescence signal (**Fig** 1.3). The signal is imaged onto an sCMOS camera for full frame image recording. 3D image stacks can be generated by either moving the sample through a stationary light sheet or moving the light sheet through the sample while synchronously refocus the detection objective.

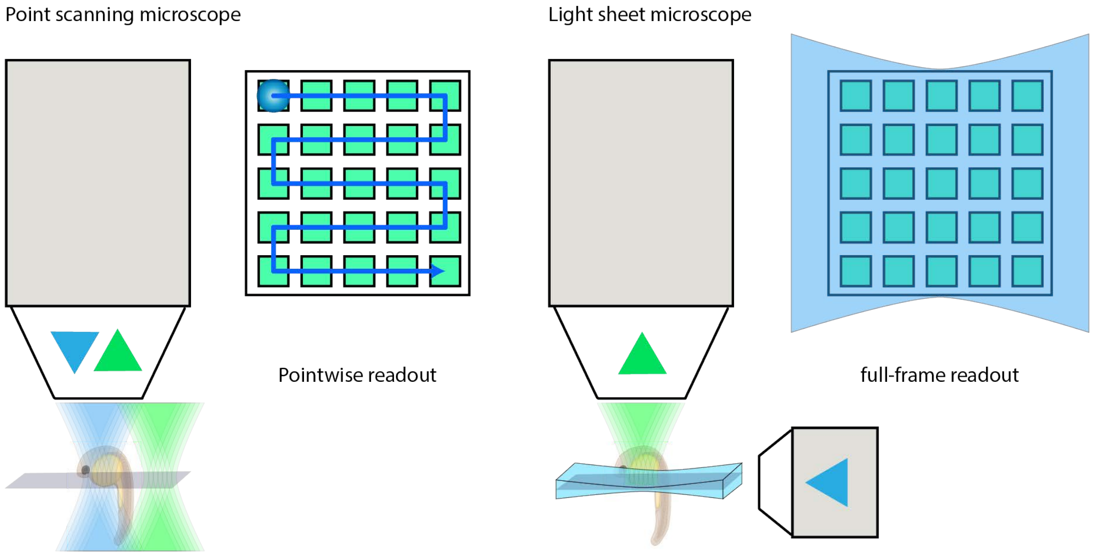


Figure 1.4 Comparison between the light sheet and confocal microscopy. Microscope that uses point scanning to generate images (such as confocal microscopy and 2-photon microscopy), the focus point has to scan through the image plane sequentially to generate an image. In SPIM, the entire plane of interested is illuminated simultaneously. At the same imaging framerate, SPIM has much higher detection efficiency.

In contrast to confocal microscopy, only the plane of interest is illuminated in the case of SPIM, and therefore the sample is exposed to much less laser light during the same 3D stack acquisition. Besides, since the fluorescence signal from the plane of illumination can be acquired simultaneously by a plane-detector, the imaging speed is much faster.

SPIM is well known for its low photo-toxicity. Since all pixels on the camera are exposed simultaneously and the entire sample plane is exposed to the laser light at the same time, the required laser intensity to achieve the same signal level in SPIM is much lower than point scanning techniques such as confocal microscopy (**Table** 1.1). As a result, SPIM can be used to perform time-lapse experiments at a much longer duration without worrying too much about sample health.

Due to its advantages, researchers have been able to utilize SPIM to generate some of the most stunning and information-rich datasets currently available. From long time-lapse of developing embryos to high-speed functional calcium imaging of neurons, SPIM has revolutionized how researchers approach 3D imaging. At the same time, SPIM is a highly versatile technology. It has been applied to imaging applications across a wide range of spatial scales. A specific variant of LSFM called Lat-

	Confocal microscopy	SPIM
Camera frame rate	50fps	50fps
Frame size	e.g. 1000×1000	1000×1000
Pixel exposure time	$\approx 20ns$	$20ms$
Laser intensity	$1000000\times$	$1\times$

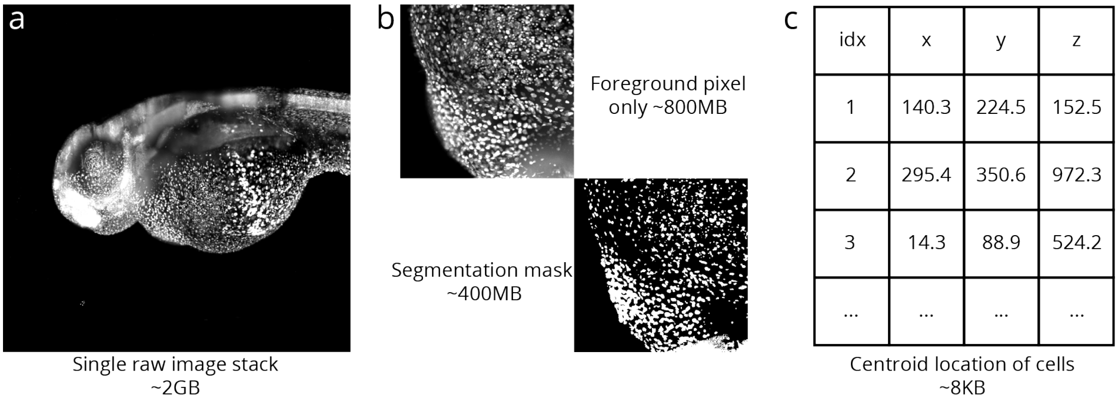
TABLE 1.1 With the same imaging speed and frame pixel count, SPIM requires significantly less laser intensity input than confocal to generate the same amount of fluorescence thanks to planar illumination and full frame readout. In practice, fluorescence intensity does not scale linearly with the excitation laser intensity due to fluorophore photo-bleaching.

tice light sheet microscopy (LLSM, B.-C. Chen et al., 2014) has been used to observe sub-cellular dynamics. At the same time, in combination with sample clearing and expansion techniques, SPIM has been used to image samples up to centimeters in size (Dodt et al., 2007). As the adaptation of SPIM as a biological discovery tool becomes widespread, SPIM was named as "Method of the Year" 2014 by the journal *Nature Methods*. However, SPIM brings not only new imaging opportunity but also unexpected challenges.

1.4 BIG DATA CHALLENGE IN LSFM

The data volume captured in a typical light sheet experiment is often orders of magnitude higher than what would be captured in a confocal experiment. The low phototoxicity of light sheet microscopy ensures a more extended experiment, and the collection efficiency usually promotes the dynamic range of images from 8 bit to 16 bit. While a confocal time-lapse session generates data in units of gigabytes (GB), a light sheet experiment, on the other hand, is usually measured in terabytes (TB). During my Ph.D. research, I have captured more than 340 TB of image data, which is way beyond the data storage capacity of a typical biology lab. It was not uncommon for institutes to adopt a light sheet system only to discover that they need to invest heavily on the storage infrastructure before being able to use the microscope at all.

In addition to requiring more space for data storage, higher data volume also means the computational power needed to visualize and analyze data is also much higher. Currently, very few open source solution can seamlessly open a typical light sheet dataset for 2D visualization (Bria et al., 2016; Pietzsch et al., 2015; Schmid et al., 2019), let alone full 3D or even 4D data exploration.



idx	x	y	z
1	140.3	224.5	152.5
2	295.4	350.6	972.3
3	14.3	88.9	524.2
...	...	...	...

FIGURE 1.5 Information conversion efficiency problem in SPIM. (a) Raw image of a zebrafish embryo with histone label at 48 hours post fertilization. Image size around 2 gigabytes per stack. (b) Region of interest containing only foreground pixels and the corresponding binary mask. (c) List of coordinates of the centroid locations for each cell.

The default solution to the data burden problem is to extend the investment into the storage capacity and utilize cluster grade computing infrastructure (McDole et al., 2018; Schmied et al., 2015) for analysis. In my opinion, these solutions, although able to successfully analyze light sheet datasets, do so in a brute force manner rather than addressing the fundamental issue.

Consider the following typical light sheet experimental goal: Imaging zebrafish gastrulation for 24 hours at 1 minutes interval and tracking every single cell to unravel the collective cellular dynamics. A single view is typically not sufficient to cover the entire zebrafish embryo. Therefore, multi-view imaging needs to be used. A total of 4 views or more is usually required. If each view consists of 500 frames of sCMOS images (2048×2048), each time point generates around 20 GB of data. The entire experiment totals around 28 terabytes of data. However, the final data output from the experiment only consists of tracking information of each cell. More than 10 trillion 16-bit integer voxel values in the raw image need to be converted to the cell location information, which contains less than 60 million floating point number. An information conversion ratio measured in data product size over raw data size of less than 0.001% is estimated.

$$\text{Information conversion ratio} = \frac{\text{Size of final data product}}{\text{Size of original raw data}} \tag{1.1}$$

The eventual goal of any bio-imaging experiment is to use the data product to reach scientific conclusions. The big data challenge in SPIM, in my opinion, is not only a data volume issue. The lack of effective methods to convert large image volumes into smaller but information rich data product can also stop biologists from fully utilizing the data they acquire.

1.5 Smart microscopy: Improving the information conversion efficiency

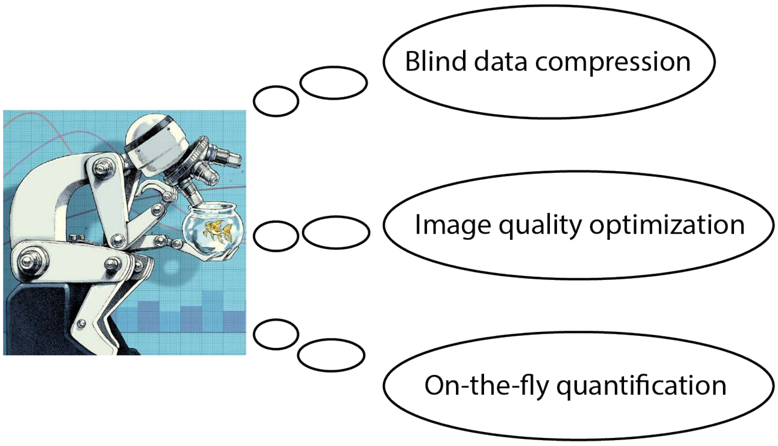


Figure 1.6 Three different approaches to improve the information conversion efficiency. Adapted from Scherf and Huisken, 2015.

In my opinion, there are two separate but related causes of the big data challenge in microscopy. Firstly, image data acquired by microscopes can have large data volume. However, not all parts of the image contains useful information needed for quantification. For example, background pixels are abundant in the raw image data but they do not contribute to the generation of the final data product. On the other hand, foreground pixels might contain varying amount of information. For example, poor image quality caused by sample scattering can significantly impact the difficulty in image analysis and impact the quality of the data product as well. Therefore, storing more data may not directly translate into a proportional increase in the amount of information that can be extracted.

On the other hand, there is a temporal disconnect between image acquisition and image processing. The data volume burden mainly occurs during the transient stage before raw images are converted into data product. Data is often stored on hard drives whose data bandwidth is much lower than the both the camera capture speed and computer memory speed. Hence there is a significant time penalty each time the data need to be written to or read from a hard drive storage. In SPIM, the time penalty can significantly prolong the entire workflow and decrease the experimental throughput. For example, transferring a 28 TB multi-view SPIM dataset from the acquisition computer to another computer for analysis through an external SSD (with read/write speed at $300MB/s$) can take up to 36 hours. In comparison, the data acquisition time is only around 24 hours. Hence it is also important to reduce the number of unnecessary storage steps between converting raw data into data products.

In machine learning, there are three main ways of achieving automation for different tasks depending on the amount of prior information available (Fully supervised, semi-supervised, and supervised machine learning approaches). Analogously, I believe there are three ways to improve the information conversion inefficiency.

Firstly, the data size can be reduced using a generalized compression algorithm. The information conversion ratio can be improved by reducing the denominator of **equation** 1.1. The amount of improvement largely depends on specific data features such as the dynamic range and SNR. This approach does not utilize much prior information related to the specific sample type and can be used in a wide variety of applications.

Another way of improving information conversion ratio is by increasing the amount of information contained in the raw data. This usually involves optical methods to improve the image quality and can, in turn, help with the downstream data processing. For example, Royer et al., 2016 have demonstrated that the image quality of SPIM can be improved by incorporating low order adaptive optics corrections to combat optical aberrations caused by the sample. In their study, image stacks acquired by the microscope are analyzed by a custom software suite during acquisition. Many optical elements (such as beam-steering mirrors) are motorized and adjusted to optimize the image quality estimated by the software on-the-fly. The method uses much prior knowledge regarding the geometrical relation between the effects of adjusting each component and imaging result. The method does not make many assumptions about the sample being imaged. The method's performance is determined by the optical quality of the sample but is principally able to be used in a variety of large field-of-view imaging applications.

The previous two methods aim to improve the information conversion ratio by either increasing the numerator or decreasing the denominator in **equation** 1.1. Alternatively, certain quantification process can be performed on-the-fly (Schmid et al., 2013). The raw data only exists in the memory buffer of the capture machine temporarily, alleviating the disk storage stress. More importantly, certain on-the-fly quantification also allows the microscope to perform more intelligently, such as not recording blank pixels unnecessarily (Long et al., 2017).

All three approaches mentioned above constitute what I believe the three pillars of a smart microscope. They share a frequent need to incorporate data processing into the image acquisition process (**Fig** 1.6). This thesis focuses primarily on building tailored light sheet based workflows that utilize one or more said approaches to realize improved information conversion efficiencies.

1.6 SCOPE OF THESIS

This thesis is divided into three parts and six different chapters.

After the **Introduction** follow the **Results**, which include four chapters. In **Chapter** 2, I compare different general data compression algorithms and their applicabilities to light sheet datasets. I also introduce some integer compression algorithms that are more efficient at compressing SPIM images. In this chapter, I also give an overview of the IT infrastructure and data pipeline used in my thesis work that is ideal for a smart SPIM workflow.

In **Chapter** 3, I introduce the smart rotation workflow, in which I incorporated the on-the-fly image analysis pipeline into a conventional multi-view light sheet experiment. The pipeline automatically decides the optimal view combination that offers the best sample coverage in an *in toto* SPIM imaging experiment. I compared the performance of a conventional multi-view imaging scheme with the performance of the new workflow. I showe that the smart rotation workflow outperforms the previous approach in terms of sample coverage with a given number of views when imaging zebrafish embryos. The smart rotation workflow is an example of improving the image quality to increase the information conversion efficiency.

In **Chapter** 4, I describe a custom workflow that I developed to quantify the phenotype of zebrafish Rohon-Beard cell axon projections. The chapter details the establishment of the entire workflow from sample preparation to image analysis. I also describe the custom image unwrapping software used in the study in detail. The software can run during image acquisition and thereby reduce the amount of data that needs to be saved. I will also introduce my on-going effort to develop the workflow further into a mid-throughput phenotype screening platform.

In **Chapter** 5, I showcase a variety of projects that I participated in during my Ph.D.

work. I introduce my contributions to building tailored imaging workflows for a wide variety of sample types. I will also demonstrate that the smart microscopy concept is key to some projects' success.

In the last part, I give a conclusion and outlook of my Ph.D. work and of the future of smart microscopy. This chapter also includes a discussion of the on-going deep learning revolution that is reaching the microscopy field and how the smart microscopy concept can benefit from artificial intelligence (AI) in the future.

All SPIM images shown in this thesis are taken with a home-built multi-view SPIM (Details in **Appendix** C.1). All software codes can be found at https://github.com/henryhetired.

Part II

RESULTS

2

Data size reduction

"Most speeches are like Texas Longhorns — a point here, a point there, and a lot of bull in between."
— *Anonymous*

2.1 Data compression for light sheet fluorescence microscopy

In order to address the data burden that is common for LSFM, the first intuitive step is to compress the image data acquired. Many different data compression methods have been tried on light sheet datasets. However, the majority of the commonly used data compression algorithms cannot be applied to light sheet datasets directly as they are too slow. Most prior works address the problem either by using dedicated hardware to accelerate the compression speed or by throwing away information in the data that is deemed irrelevant to the data integrity to achieve a better result. Here I compare some prior work on light sheet data compression and introduces a few methods that are more suitable for real-time data compression.

2.1.1 Overview of prior work

2.1.1.1 Lossless and lossy compression

Data compression, in essence, converts the raw data format into a format with a smaller volume that can be returned to its original format when needed. If the data can be recovered to be exactly the original raw content, then the compression is called **lossless**. Alternatively, the compression process can also throw away some details that are not significant for the overall information integrity of the data. In this scenario, the compression is considered **lossy**. Generally, lossy compression has a better compression ratio than lossless compression.

Numerous best performing general purpose compression algorithms have been used to compress light sheet data. Amat et al., 2015 use a parallel version of *Bzip2* to losslessly compress data for storage in HDF5 format. The compression ratio can be vastly improved if a masking step is performed where background pixels are discarded. Balázs et al., 2017 use a modified version of the well-known JPEG compression algorithm to compress image data in a lossy manner. In order to speed up the data compression processing, Balázs et al., 2017 utilized graphics processing units (GPU) to achieve a compression rate similar to the camera acquisition speed. Similar to the method mentioned above, data can be discarded in favor of the compression ratio. It is worth noticing how Balázs et al., 2017 justify the use of lossy compression in a scientific workflow: They compared the downstream processing results using the raw data and the lossy compressed data and showed that it is possible to tune the parameters such that the lossy compression does not affect downstream analysis result. They termed the method as 'Noise-level' compression. In a real experiment, it is not clear how one could know *a priori* how the parameters can be set before running the analysis pipeline once. Also, the use case is limited to imaging settings where the foreground pixels have a significantly different brightness level compared to the background noise. It is not clear how the method would perform when the signal to noise level is low. In my opinion, the method is most useful when the data acquisition parameters do not change between experiments. Therefore, a configuration can be optimized and kept consistent between experiments.

2.1.1.2 Content aware data compression

Instead of compressing data blindly, there are also methods to reduce the data volume by taking into account the specific spatial features of light sheet datasets. Schmid et al., 2013 notice that during zebrafish embryonic gastrulation, the cells that were being imaged lies on the yolk. Hence the useful information lies on a spherical shell. By fitting a sphere to the image volume and only store data on the shell, a 1% data compression ratio can be achieved. The method is very application specific and cannot easily be modified to fit other experimental settings.

Cheeseman et al., 2018 describe the adaptive particle representation (APR) that pragmatically varies the effective pixel size such that background voxels are given less storage weight compared to foreground voxels to achieve data compression. APR requires only a single setting on the noise level, so it applies to virtually all sample types. The compression ratio that can be achieved by APR is sample and experiment dependent. Similarly to the sparse matrix representation, the compression ratio is dependent on the density of foreground voxels within the dataset.

Both methods above are lossy as background signal that was discarded cannot be recovered. It is worth noting that the method described in Schmid et al., 2013 allows the image to be transformed from 3D into 2D via spherical projection methods. The transformed image are both smaller in size and easier to analyze. The image compression step also enables new avenues for data analysis. On the other hand, the APR is a fundamentally new way of representing the image data. Any existing algorithms that run on raster pixel data can be adapted into utilizing the APR data structure. APR reduces both the storage volume of the dataset and the amount of memory needed for processing tasks. However, as of the time of writing this thesis, there is not a large suite of commonly used image processing algorithms that are implemented in APR.

2.1.2 General purpose compression methods

In order to incorporate the data reduction process into a smart microscopy workflow, I first tested some general file compression methods. The three main criteria for my evaluations are:

- **Compression rate**: The volume of the compressed data relative to the raw data.
- **Compression speed**: The time it takes to compress a single frame of light sheet data.
- **Decompression speed**: The time it takes to decompress a single frame of light sheet data.

Compression rate is conventionally the most important performance metric for data

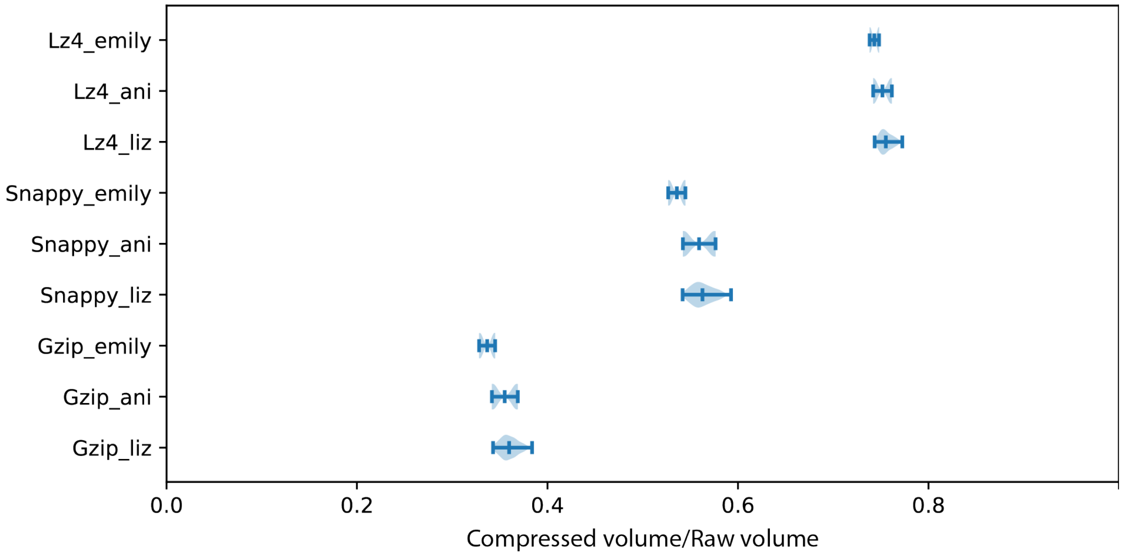


FIGURE 2.1 The compression rate of some general compression methods on three separate light sheet datasets. Details of the three datasets are described in **Appendix** A. Only performances of *Lz4*, *Snappy* and *Gzip* are shown since other compression methods are significantly slower.

compression applications. Many best performing compression algorithms such as *Bzip2* sacrifice compression speed in favor of compression rate. The **compression speed** of these algorithms are too slow for on-the-fly applications without using specialized hardware such as GPU or FPGA for parallelization and acceleration. **Decompression speed** is also an important factor as it defines the ease-of-access of the compressed data for the end users. All testings performed in this section use a single thread of an *Intel* Xeon E5-2698 v4 CPU. The restriction is to ensure the compression algorithm does not require dedicated computing hardware and can be run on any computer.

To evaluate the performance of some of the most popular general compression algorithms, I uses three representative light sheet datasets that I acquired during the my thesis study. The datasets have a varying amount of SNR and number of foreground pixels per frame (See Appendix A for a detailed description of the dataset). Each dataset contains 100 image stacks taken with a home-built multi-directional SPIM system. The image dimension is 2048 × 2048. Each image stack is compressed ten times to account for the performance fluctuation of the CPU. Only lossless compression algorithms are compared to ensure a parameter-free pipeline and compatibility with fully automated on-the-fly usage.

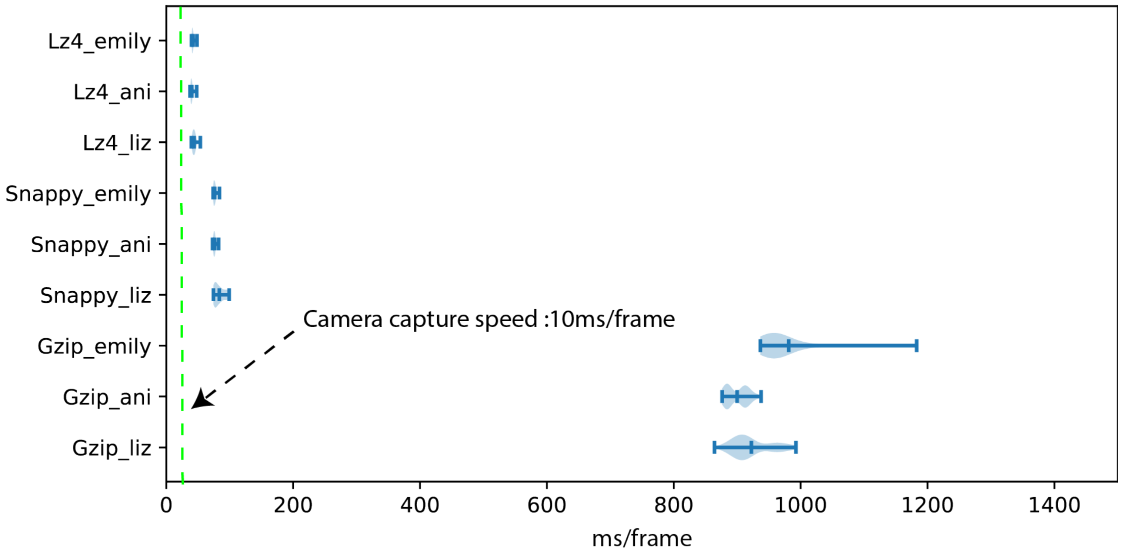


FIGURE 2.2 The compression speed of some general compression methods on three separate light sheet datasets.

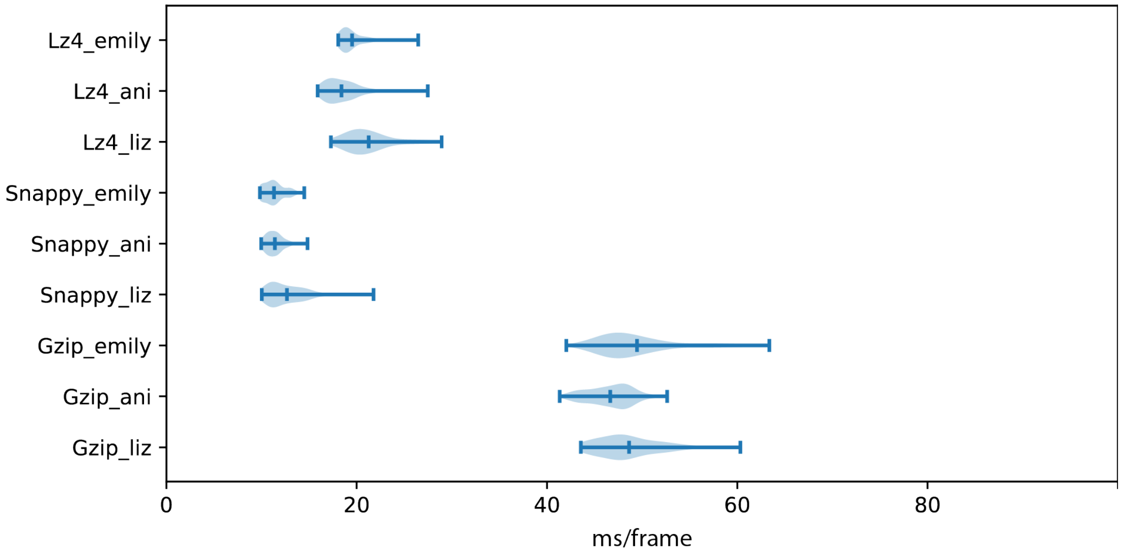


FIGURE 2.3 The decompression speed of some general compression methods on three separate light sheet datasets.

2.1.2.1 Results

Most commonly used file compression algorithms such as *Bzip2* (Seward, 1996) and *LZO* (Oberhumer, 1996) are primarily optimized for compression ratio rather than speed. Their compression speed is usually an order of magnitude lower than the image acquisition speed, which makes them incompatible to on-the-fly applications without acceleration. In my opinion, these slower but more efficient compression methods are more suitable for transferring small datasets. As the transporting media (external hard drive or network) has limited bandwidth, it is important to reduce the data size as much as possible. Most methods compared here are built into most commonly used operating systems such as *Linux* and *Windows* so they are very accessible to users.

The three top-performing compression algorithms that I tested show trade-offs between compression rate and compression speed (**Fig** 2.1,2.2). *LZ4* (Collet, 2011) is the fastest in terms of compression speed but only reduce the data volume by around 20% whereas the slowest algorithm *Gzip* reduce the data volume by up to 60%. *Snappy* (Google, 2011) has the best overall performances. However, the compression speed is still not comparable to the camera frame rate (**Fig** 2.2).

It is worth noting that specific implementation of these compression methods are not entirely suitable for 16-bit integer data. Since these algorithms are supposed to be general purpose, they would consider any file as an integer array stored in the native system bit-rate (32 or 64). Therefore, I had to manually cast the image data into higher bit-width before compression. As a result, the data compression process uses higher amount of system memory and processing power.

2.1.3 Lossless integer compression for smart microscopy

As general compression methods are not fast enough for real-time applications, I set out to study how to compress data more efficiently. General compression methods do not make any assumptions on the nature of the data. To ensure overall applicability, the methods studied above utilize multiple compression methods in conjunction. For example, *Bzip* utilizes compression schemes such as Huffman encoding (Huffman, 1952) and run-length encoding (Robinson and Cherry, 1967) in sequence. Different steps contribute differently to the overall performance, and their respective contributions depend on the data type. Therefore it is critical to find the main contributor to light sheet datasets' data bloat and thereby using only the necessary compression step.

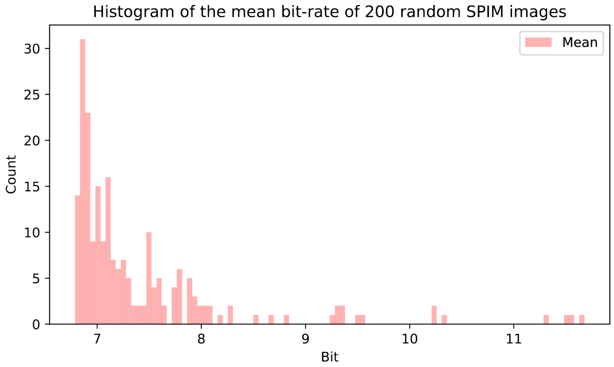


Figure 2.4 Histogram of the mean bit-rate of 200 different SPIM image stacks. The data were randomly selected from all the image data I acquired during my Ph.D work.

2.1.3.1 ADAPTIVE BITRATE FOR DATA COMPRESSION

The most commonly used cameras in LSFM such as *Andor Zyla 4.2 Plus* and *Hamamatsu Orca Flash 4.2V3* output 16-bit unsigned integer data at up to 100 frames per second. The data container can, therefore, store pixel value ranging from 0 to $2^{16} - 1$. It is worth noting that the two primary imaging applications for SPIM are high-speed and long-term imaging. In the case of high-speed light sheet imaging, the camera exposure time is short, and as a result, the signal is usually dim. When performing long-term 3D imaging with SPIM, the laser power is also set conservatively to ensure consistent sample health over a long period, the resultant images also usually do not saturate most pixels. As a result, many light sheet data does not span the full range that the data container offers. Contrast stretching and histogram equalization are often needed before the image can be visualized. In my experience, the average bit-rate of most datasets is significantly less than 16 (**Fig** 2.4). I postulate that a significant amount of data reduction can be achieved by reducing the data bit-rate adaptively (**Fig** 2.5).

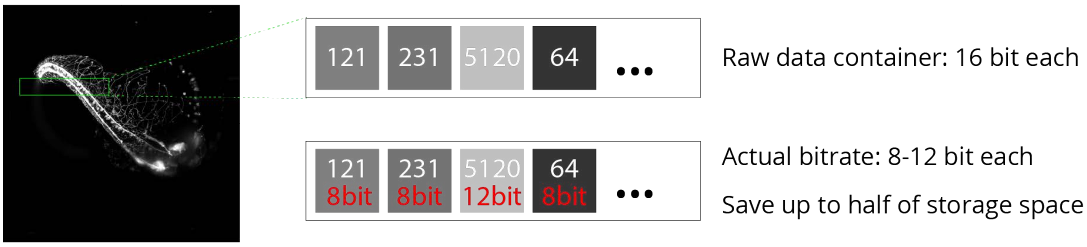


Figure 2.5 Principle of variable byte compression. A fixed bitrate for all pixels is wasteful as different pixel value spans different value range.

2.1.3.2 Results

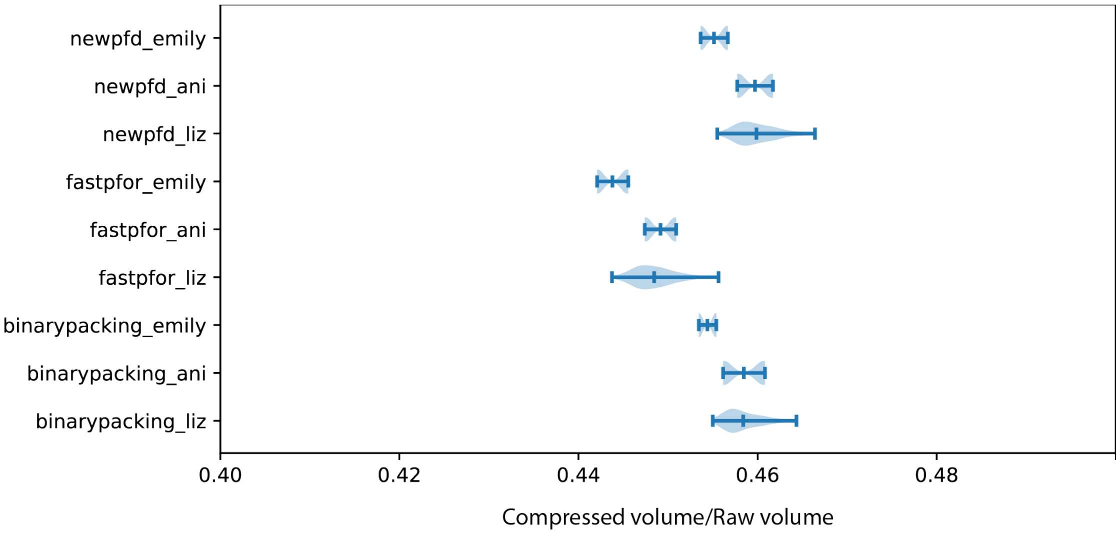


Figure 2.6 Compression rate benchmark of integer array compression methods. The compression rate of three integer compression methods on three separate light sheet datasets.

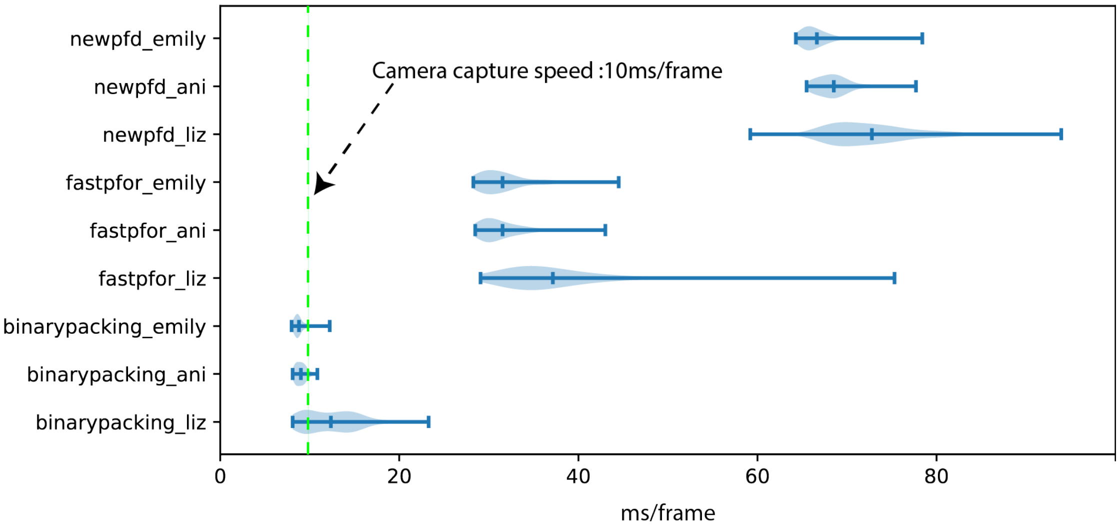


Figure 2.7 Compression speed benchmark of integer array compression methods. The compression speed of three integer compression methods on three separate light sheet datasets.

I tested a variety of state-of-art lossless integer array compression algorithms that utilize block-based variable bit-rate on real light sheet datasets (**Fig** 2.6,2.7,2.8). Since the compression performances largely depend on the actual bit-rate of the dataset, the performance is a lot more consistent compared to the general compression methods.

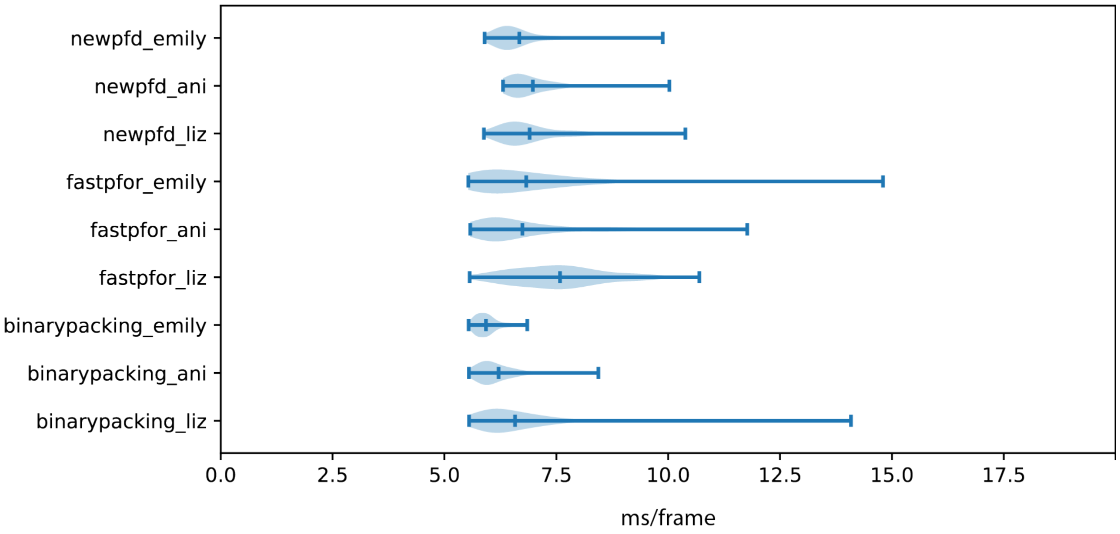


FIGURE 2.8 Decompression speed benchmark of integer array compression methods. The decompression speed of three integer compression methods on three separate light sheet datasets.

The image dataset volume can be consistently reduced by half (**Fig** 2.6). This is consistent with the hypothesis that the main contributor to the data bloat in SPIM datasets comes from low bit-width utilization. Therefore the compression rate depends only on the real data bit-rate within the sample. The compression rate is comparable if not better than general algorithms when applied to real light sheet datasets. Moreover, the compression speed of these integer compression algorithms is comparable to the image acquisition speed, with *binary packing* method being the fastest (**Fig** 2.7). Specifically, the *binary packing* compression method compresses data at around 100 frames per second and decompresses at around 150 frames per second, making it the overall method of choice for light sheet data compression. (Lemire and Boytsov, 2012).

2.1.4 Discussion and outlooks

In this section, I compared different compression methods that have been applied to SPIM datasets to a more efficient class of compression algorithms. Integer compression methods are much faster and do not have any sacrifices in compression ratio, paving the way to real-time image compression in SPIM without dedicated hardware. It is worth noting that the integer compression algorithms tested here do have one tunable parameter which was left as default (64 integers per block) in the testing above. The encoding window defines the expected consistency in bit-rate within the data and can be tuned to the foreground pixel density. I have incorporated the compression step into my custom SPIM imaging workflow. The software currently only has a command line interface. I plan to build the compression method into the Fiji image reading process to make the method more accessible to other users.

2.2 Workflow framework for smart microscopy

In order to incorporate the on-the-fly image analysis step into the imaging process, there are many practical considerations regarding data handling that have to be addressed first. Commonly used data pipelines have many bottlenecks that are especially prevalent in the presence of large SPIM datasets. Eliminating these experiment-stopping bottlenecks is critical for the smooth operation of an integrated image analysis pipeline.

2.2.1 Conventional data pipeline in microscopy

There are usually three locations where image data are stored during an imaging experiment. At the data source, images captured by the microscope camera are directly written to a computer that is used to control the microscope. Users then need to copy the data away from the microscope machine either using an external hard drive or over the network. The data is usually copied to a powerful workstation that has either commercial or custom made image analysis software installed. The workstation is usually also used as a data storage server where the entire laboratory's image data are kept.

The workflow was sufficient for image modality such as confocal microscopy that does not generate a considerable amount of data. The amount of time required to move data from one location to the next is small compared to the other steps of the workflow. However, LSFM is capable of generating about 800 MB per seconds of data. The data throughput is often pushing the limit of the hard drives' read speed. As a result, copying data away from the microscope computer becomes a time-consuming step. It is common for a time-lapse SPIM experiment to generate 10s of terabytes of data and to transfer the data using an external hard drive can often take days. As the hard drive bandwidth on the capture computer is not sufficient to enable simultaneous read and write at high speed, the microscope has to remain idle during data copying. The overall duty cycle of a microscope is, therefore, significantly reduced. Other than the data transfer bottleneck, a single image analysis workstation is also inefficient in a multi-user environment. Image analysis steps in SPIM often take up significant

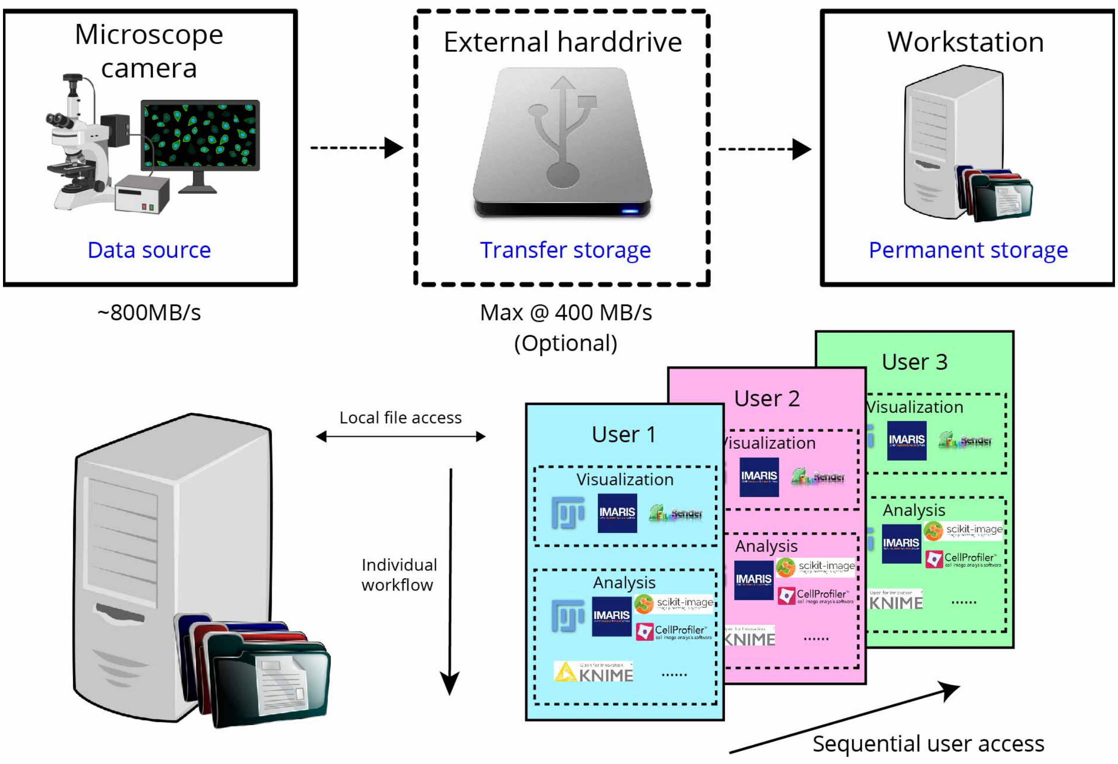


Figure 2.9 The conventional data pipeline in a microscopy workflow. The need to relocate the data from the microscope capture machine to the final storage potentially limit the experimental throughput. The analysis workstation also has sequential user access, creating an additional bottleneck.

computing resources and time. Sequential user access often results in long user wait time while still having a long machine idle time when the previous task was completed without notifying the subsequent user (**Fig** 2.9).

Schmied et al., 2015 and Daetwyler et al., 2019 described software workflows to perform the data transfer and processing automatically. This reduces the amount of manual labor required but does not address the inefficiency issue. The need to transfer data between storage spaces in the first place is the root cause. Also, the pipeline presented by Daetwyler et al. is very application specific. The pipeline reported by Schmied et al. streamlined the data processing steps and can utilize parallel computing infrastructure. However, the solution is not easily scalable for labs that do not have access to large computing clusters. More importantly, the pipeline does not address the multi-user conundrum. Here I introduce an integrated data pipeline that addresses both problems and also enables on-the-fly imaging process for smart microscopy applications.

2.2.2 Integrated data pipeline for smart microscopy

In order to eliminate the need for temporary transfer storage solution, I have tested a few methods to move the data directly to the permanent storage solution. The most straightforward solution is to write the data directly to a portable storage such as an external hard drive. The analysis workstation can access the data by simply attaching the portable hard drive into the machine. However the capacity of the external hard drive is usually limited to below 10TB. It is not suitable for long SPIM timelapse experiments. In addition, single external hard drives does not have any redundancy protection. If the external hard drive becomes faulty, all the data on the hard drive becomes lost. There are also small hard drive enclosures that can house multiple drives. They also have RAID support to provide the necessary redundancy for backup. However, they are much less portable. In collaboration with the IT team at *Morgridge Institute for Research*, we tested a variety of network attached storage (NAS) systems. We built a centralized storage server that is accessible by all the microscope computer via a high-speed network. The fileserver is set up to have RAID 10 configuration that has redundancy protection. Data generated by all the microscopes are streamed directly to the fileserver and never reside on any temporary storage media. We found the solution to be both robust and highly scalable. Different labs can adopt the same system and choose the storage capacity and network speed based on their specific throughput.

Another benefit of the centralizing the storage is that we enabled parallel data accessibility to users in the same lab. Different users can mount the fileserver as a remote hard drive on any computer and access the data at the same time. We also attached a powerful *Linux* workstation to the fileserver via 40Gbps network. Users can connect to remote sessions on the workstation and perform more demanding image analysis tasks. With the centralized storage server and remote accessible analysis machine, we addressed both the data transfer and user access bottlenecks (**Fig** 2.10).

Hardware specification for the data pipeline used in the Huisken lab at *Morgridge Institute for Research* can be found in **Appendix** C.2.

Although similar microscopy data management pipelines exist, they are optimized for smaller datasets. One of the most popular integrated microscopy data platform, *Open Microscopy Environment* (OMERO), has a similar backend setting. However, the frontend web-based image visualization and analysis interface is not ideal for large SPIM images. As a result, users have to use other software that is separate from OMERO, complicating the workflow. Also, OMERO currently does not offer any hardware integration such that the microscopes can be directly accessed through OMERO.

In addition to addressing the bottlenecks of conventional workflows, our new integrated pipeline is also critical for any on-the-fly image analysis tasks. The network communication between the microscope computer and the data server as well as the analysis server are bidirectional. Raw images are directly accessible by the analysis machine and therefore, can be analyzed as soon as possible. A bottleneck-free pipeline is particularly useful for applications where a direct feedback loop between the hardware control computer and an image analysis engine is needed. For example, adaptive optics applications require the microscope to react to the image quality assessment to perform reconfiguration. Any delays in the pipeline would increase the reaction time and may result in the optical correction no longer valid for the rapidly changing biological tissue.

In my implementation, I use the **TCP/IP** protocol to pass messages between the microscope and the analysis engine. The microscope message usually contains an analysis request, such as data compression on the file server. The analysis engine then can send either processing status or specific imaging instructions back to the microscope.
It is worth noting that all parts of the data pipeline can be scaled up or down, depending on the specific needs of the laboratory. I have implemented a command-listener software that can be run on the analysis engine (Huiskenlab_commandprocessor, repo). The command processor can receive and manage processing request and run custom *Java* and *Python* scripts as well as *Fiji* macro scripts in a multi-threaded manner. I am currently working on extending the software to allow multiple processing queues managed separately based on user identity. The command processor and the integrated data pipeline are both utilized in a variety of smart microscopy applications. (see **Chapter** 3 and 4).

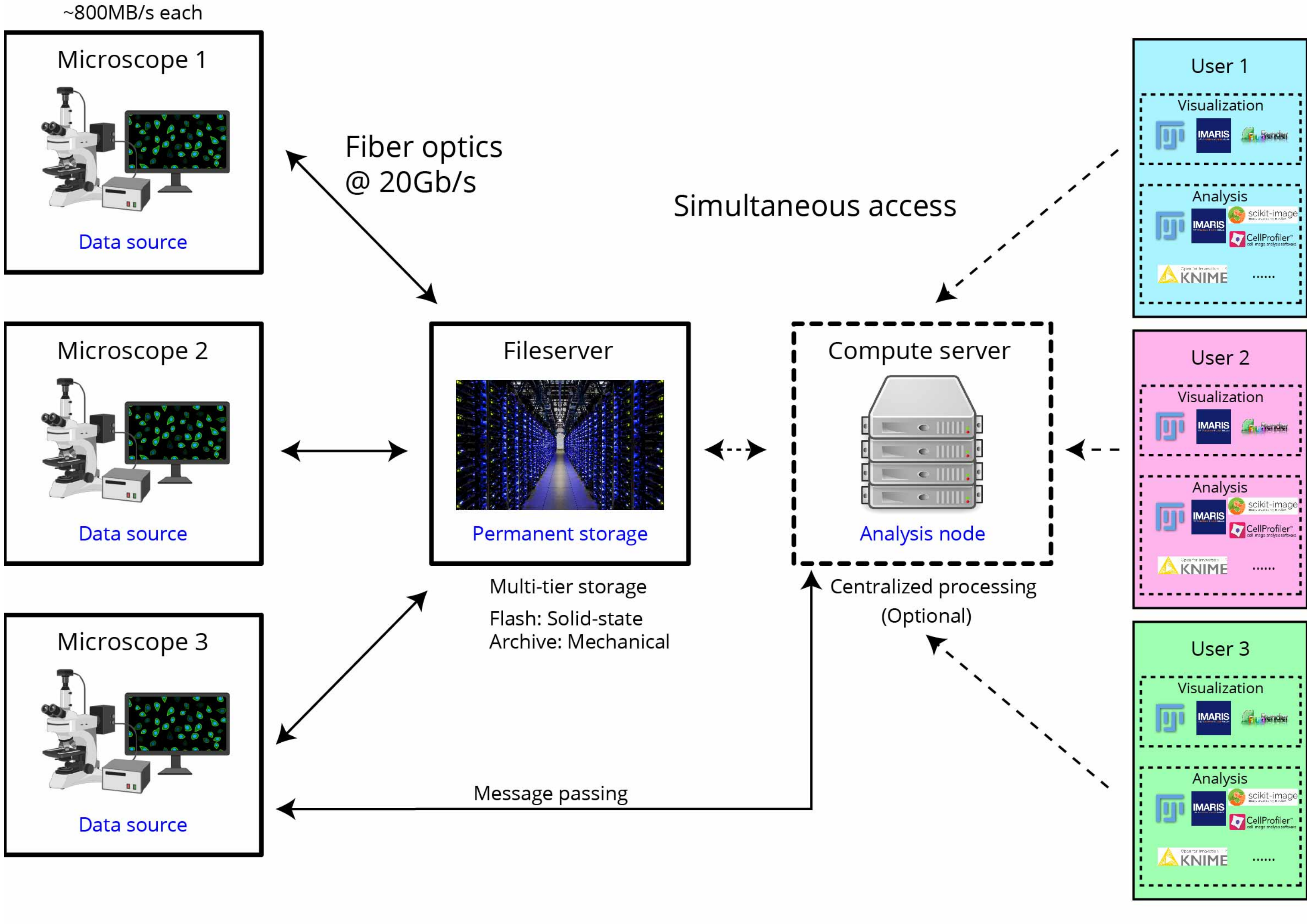


FIGURE 2.10 The integrated data pipeline removes data access and data transfer bottlenecks. The pipeline also allows direct communication between the microscope and an analysis machine, which is critical for smart microscopy applications.

3

Smart rotation workflow to improve light sheet sample coverage

"The less there is to justify a traditional custom, the harder it is to get rid of it."
— Mark Twain, The Adventures of Tom Sawyer

3.1 Evaluation of sample coverage in multi-view SPIM

Many *in vivo* applications require a three-dimensional field-of-view of more than 500 µm^3. Particularly in such thick samples, aberrations caused by inhomogeneous optical properties can severely degrade the microscope's performance (**Fig** 3.1). Most common types of aberration include absorption, refraction, and scattering. Photon absorption by biological tissue (such as pigments) can cause shadowing effects that become especially apparent in SPIM's orthogonal illumination. Light refraction due to the inhomogeneous refractive index can redirect the path of photons, resulting in a range of optical artefacts including defocus and excitation beam steering. Photon scattering by complex biological tissue can broaden the light sheet and decrease the achievable resolution of the optical system. The combination of the effects mentioned above limits the actual penetration depth of excitation photons. The same degradation also applies to fluorescence photons, resulting in a weaker signal and lower than expected detection resolution. The realized imaging depth in a SPIM system is therefore sample and experiment dependent. Imaging depth is usually constrained to around 200 to 300 µm in relatively transparent samples such as zebrafish, not sufficiently deep for large FOV applications such as *in toto* development imaging.

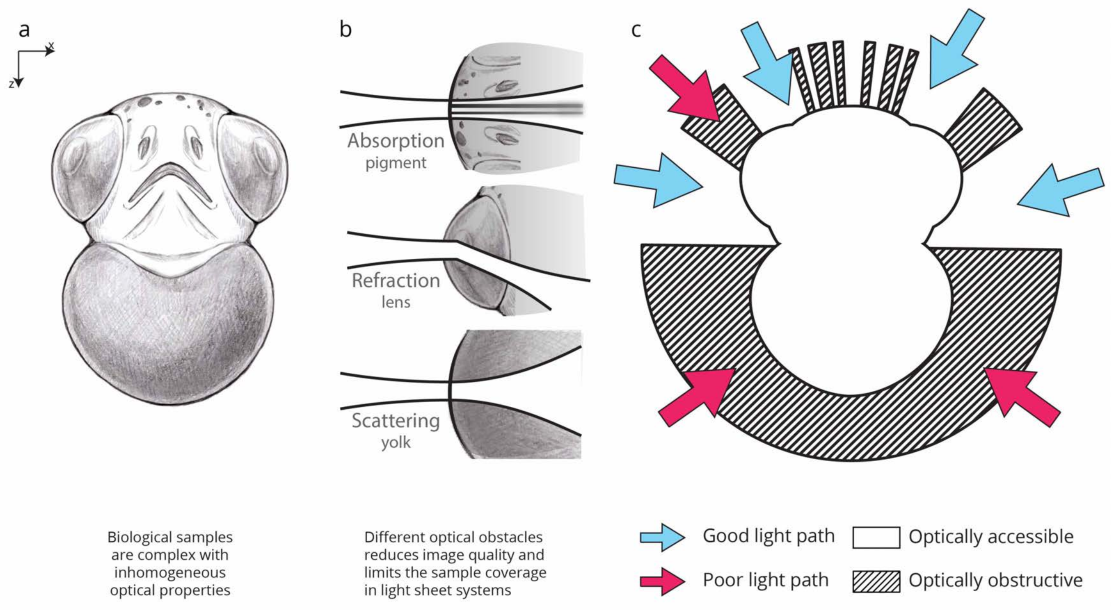


FIGURE 3.1 (a) Illustration of a zebrafish embryo in anterior-posterior view. (b) Illustrations of common optical artefacts induced by complex biological samples. (c) Different views result in different image quality due to sample inhomogeneity.

3.1.1 MULTI-VIEW IMAGING TO IMPROVE SAMPLE COVERAGE IN SPIM

Multi-view imaging in light sheet fluorescence microscopy is widely used to increase the sample coverage for *in toto* applications. By taking additional images of the same sample from a different viewpoint and merging these data, the overall image coverage is improved. Additional views can be obtained by rotating the sample and thereby changing the relative orientation of the sample to the imaging objectives (Schmid et al., 2013; Swoger et al., 2007). Alternatively, the optical paths can be modified to incorporate additional objectives around the sample. By placing an additional illumination objective opposing the first one, illumination coverage can be improved (Huisken and Stainier, 2007). Similarly, by adding another detection objective opposing the first one, the signal from the sample's far side (relative to the first objective) can be collected with higher efficiency (Krzic et al., 2012; Schmid et al., 2013; Tomer et al., 2012). Moreover, all objectives can be used for both illumination and detection, giving a total number of 8 views (Chhetri et al., 2015). In comparison, sample rotation typically requires no modifications to the optical system; simply a rotational stage is needed to turn the sample in the microscope. Multi-objective SPIM is usually faster as multiple cameras can be aligned so that their images are registered inherently without the need for post-processing, while multi-view datasets taken by sample rotation usually require an additional registration step. Otherwise, sample rotation does not substantially increase the overall imaging time with modern, fast stages. Most importantly, sample rotation gives the maximum degree of flexibility in picking the ideal orientation. The two methods have been used in conjunction to efficiently achieve complete coverage of early zebrafish embryos (Schmid et al., 2013).

A priori knowledge of the sample's features that may be optically obstructive and need to be avoided, such as eyes, dense fatty tissue, and pigmentation, can assist in selecting better angles for imaging. However, in a typical multi-view light sheet experiment, the user defines the angular views manually and often blindly or only based on a qualitative inspection of low-resolution images at the beginning of the experiment. The views are usually equally spaced and predefined in number based on the temporal resolution requirement of the experiment and the speed of the camera. Hence, the angle selection process is highly subjective, leading to inconsistent results between imaging experiments. In an attempt to achieve good sample coverage, users often end up acquiring more views than necessary. Unfortunately, phototoxicity scales linearly with the number of views, and excessive exposure of the sample to laser light can lead to bleaching and severe detriments to sample health.

In preparation for a time-lapse experiment, users typically cannot account for the natural changes of the sample over time. Here, the angular view configuration typically stays the same, although the sample's development affects its optical properties at the same time. Hence, even if the user can manually select a good set of views at one time point, these views may easily become sub-optimal during a time-lapse, yielding unsatisfactory results at later stages. Therefore, a workflow where imaging views are adaptable to sample changes is needed.

3.1.2 Method to evaluate sample coverage in a multi-view SPIM system

To be able to dynamically choose the imaging views during acquisition, an efficient and robust way of evaluating the sample coverage is needed.

In standard 2-lens SPIM (one illumination and one detection objective), the optical coverage is limited by optical obstructions in both the illumination and detection path. Only a small angular portion of the sample that is relatively close to both can be imaged well. Hence, having more views surrounding the sample would offer a more complete coverage (**Fig** 3.2). Views can be added either through placing additional objectives around the sample or rotating the sample around the vertical axis (**Fig** 3.4).

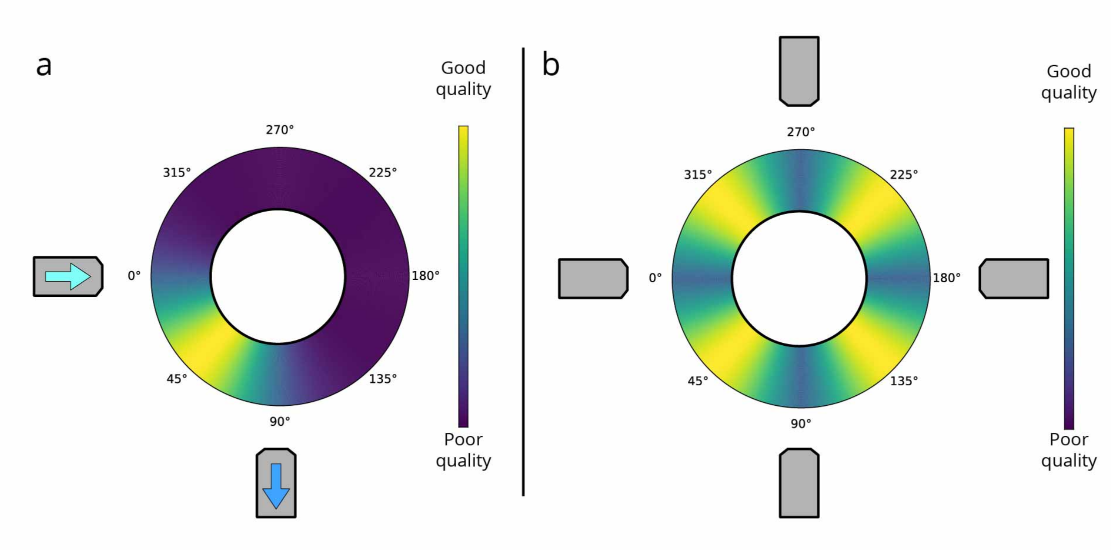


Figure 3.2 Assumption on how multi-view imaging improves sample coverage. (a) A single view only covers part of the sample's angular range. (b) Additional views provide increased angular coverage in a linear fashion.

The hypothesis above assumes that different angular regions exhibit the same imaging response when illuminated at the same relative angle. Therefore, the angular regions that are far away from both the illumination and detection angle are expected to give worse imaging result. The assumption only holds if the sample has uniform optical properties and labeling density. However, in biological samples neither the refractive index nor the fluorophore distribution are spatially homogeneous (**Fig** 3.3). Hence, blindly applying the assumption to a real biological sample can result in lower information gain than expected when using sub-optimal angles in a multi-view acquisition. There were previous methods that describe evaluation of the quality of SPIM images (Huaman et al., 2018; Royer et al., 2016). However, the image quality map of an image does not quantify the sample coverage. Image quality is often estimated on a pixel by pixel basis from a single image. The image quality of the same physical area might become completely different if imaged from a different view. Multiple image stacks from different views need to be compared to offer insights on how optimal image quality can be achieved.

3.1.2.1 Formulation

Here we introduce a formulation to evaluate the imaging coverage C_α of an angular region α. If the sample coverage is measured in terms of number of foreground voxels,

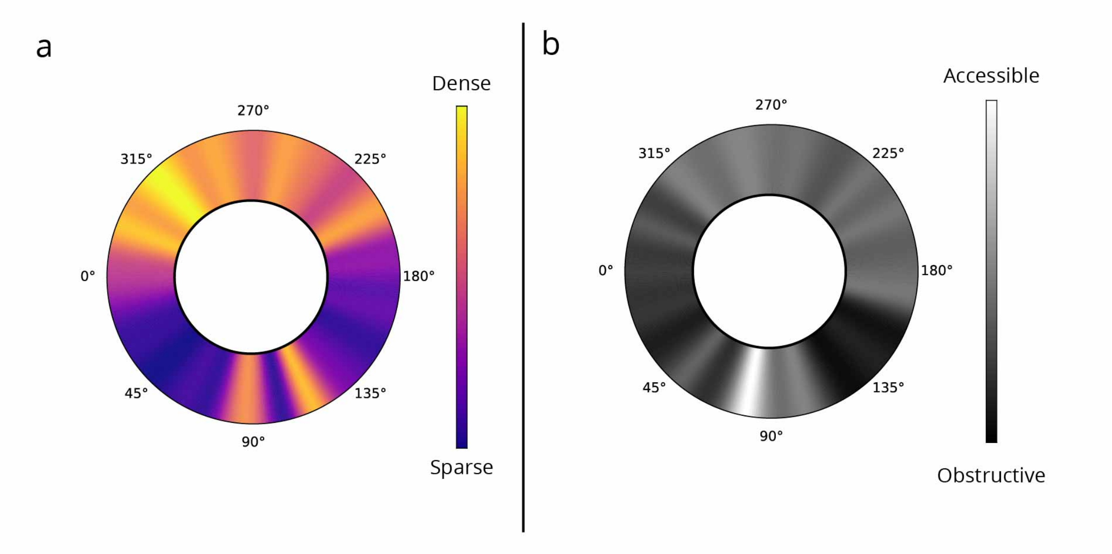


FIGURE 3.3 Angular fluorophore distribution and optical accessibility distribution. (a) Angular fluorophore distribution of a 2 day old zebrafish embryo with nucleus labelling. (b) Optical accessibility of the same embryo

then C_α could be estimated as a von Mises distribution:

$$C_\alpha = \frac{A}{2\pi I_0(\kappa)} e^{\kappa cos(x-\mu)} \tag{3.1}$$

Where x is the imaging angle and μ is the imaging angle where the maximum image coverage for this angular region is achieved. I_0 is the zeroth order Bessel function of the first kind. κ measures the concentration of the distribution, which reflects the optical accessibility of the angular region. The more concentrated the distribution in α is, the less likely a randomly selected imaging view would provide a good result. The amplitude A, encodes the underlying angular fluorophore distribution. In the conventional multi-view assumption, A and κ stay constant, whereas μ is always centered between the illumination and detection paths. In this model, increasing the number of (equally spaced) imaging views results in a linear increase in imaging coverage (**Fig** 3.2). In real biological samples, however, angular inhomogeneity in fluorophore distribution and optical accessibility introduce additional complexities in finding the optimal view configuration, resulting in an image with much less content than predicted (**Fig** 3.3).

I quantify the imaging coverage C_α with the following approximation:

$$C_\alpha \approxeq N_{\text{foreground}} \tag{3.2}$$

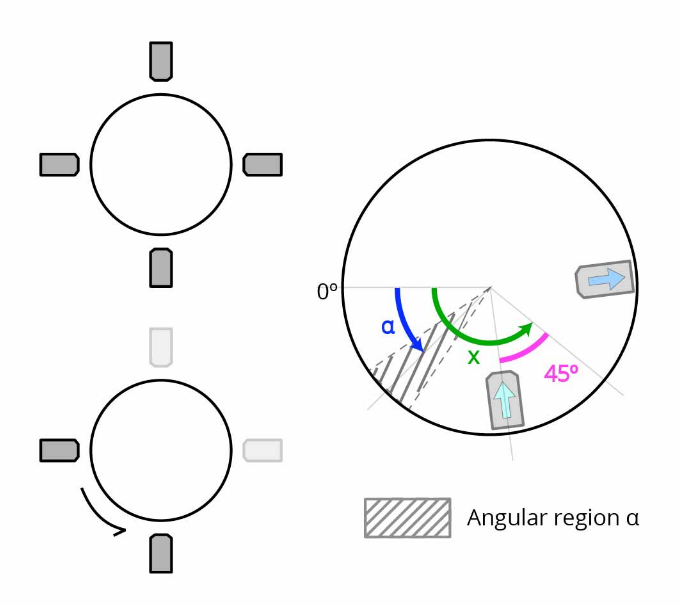


FIGURE 3.4 Illustration of angle definitions in sample coverage evaluation

Where $N_{\text{foreground}}$ is the number of foreground voxels in the image stack. The image is first split into equal sized blocks and each block is transformed using discrete cosine transform (DCT-II):

$$F_{dct}(u,v) = \frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}\delta(i)\delta(j)cos\left[\frac{\pi u}{2N}(2i+1)\right]cos\left[\frac{\pi v}{2N}(2j+1)\right]I(i,j) \tag{3.3}$$

where $I(i,j)$ is the image intensity value at pixel location (i,j) and:

$$\delta(t) = \begin{cases} \frac{1}{\sqrt{2}} & \forall t = 0 \\ 1 & \forall t \neq 0 \end{cases} \tag{3.4}$$

The spectral entropy of the transformed image is then calculated (Royer et al., 2016; Shannon, 1948):

$$S_{\text{Shannon}} = \sum_{i=1}^{n}\sum_{i=j=1}^{n} p_{i,j}\ln p_{i,j} \tag{3.5}$$

where

$$p_{i,j} = \frac{P_{i,j}}{\sum_{i,j} p_{i,j}} \tag{3.6}$$

$$P_{i,j} = \frac{F_{\mathrm{dct}}(i,j)^2}{N^2} \tag{3.7}$$

where N is the size of the patch, usually set as a power of 2 for easier GPU processing.

The resultant entropy image is then thresholded based on a predefined background level. The number of foreground pixels in the entropy image can then be used as a comparative measure of information content.

Although *Royer et al.* also use discrete cosine transform and Shannon entropy to evaluate the image, they believe the entropy image itself represents image quality. Therefore optimizing the entropy through adjusting optical components can improve image quality. I argue that it is not possible to distinguish between poor image quality and a lack of fluorophore in itself using entropy image alone. Therefore it is foreseeable that in cases where labeling is sparse, minimizing the entropy image may result in the microscope searching for the more densely labeled region instead of improving image quality. In my workflow, I use the entropy image as a measure of information content, which is affected by both the labeling density of the sample and optical performance of the microscope. It is important to isolate the contributions from both components.

3.1.2.2 Results

To verify the validity of our image coverage measurements, I applied our method to a living animal embryo. I imaged a zebrafish embryo (*Tg(h2afva:h2afva-mCherry)*,48hpf), in which all nuclei were labelled, from 24 different angular views for evaluation. 24 views were chosen since the amount of time required to take 24 image stacks roughly equal to the typical interval used in a standard time-lapse experiment.

As expected, the angular area of the image that is well-imaged sat between the illumination and the detection objective and rotated with the sample (**Fig** 3.7). The evaluation metric correctly identified well-imaged areas. Data quality degraded significantly as imaging depth increased along with both the illumination and detection direction. However, the fluorophore distribution was highly heterogeneous and optical accessibility spatially varied, resulting in a complex contour of well-imaged areas in each view (**Fig** 3.5 a-c).

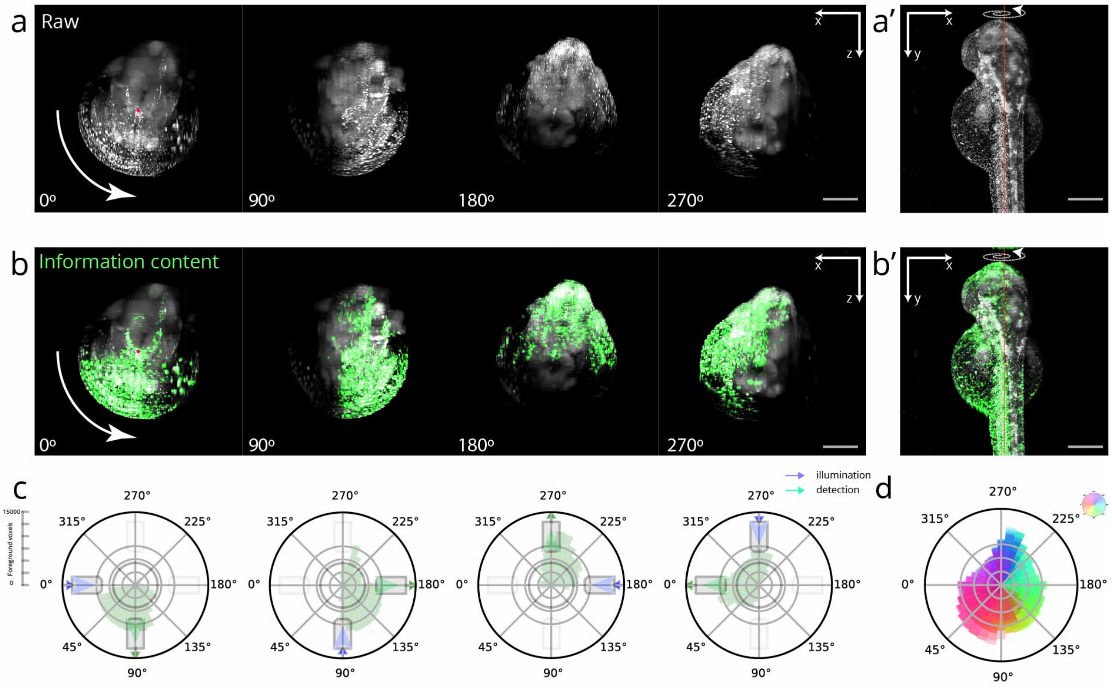


FIGURE 3.5 Sample coverage evaluation of a zebrafish embryo. (a)Anterior-posterior (x-z) view of the zebrafish embryo (*Tg(h2afva:h2afva-mCherry)*) at different angles imaged at 48 hpf. Arrow indicates the rotation direction. (a′) Dorsal view (x-y) of the zebrafish embryo. (b,b′) Information content (green) overlay with raw data (gray). Contrast adjusted with the same threshold between views.(c) Information content summarized by angle of different views measured by C_α. (d) Overlaying angular information content of 24 different views, color-coded by angle. Scale bar length represents 150 µm

The equally spaced views gave a varying amount of information and often overlapped in providing good coverage for a specific angular region (**Fig** 3.5 d). We then summed the number of foreground voxels by angular region and created a sample response profile. The response curves for each angular region were then fitted to **equation** 3.2 (**Fig** 3.6 a). It became evident that different regions had vastly different responses to the imaging angle. The coverage varies with the imaging angle. The underlying labeling density dictates the maximum amount of information each angular region contains. If a blind imaging approach is employed, where angles are equally spatially spaced, the overall coverage can be lower than expected due to inhomogeneity. Using a more flexible angular spacing can improve sample coverage (**Fig** 3.6 b). If the fluorophore distribution and optical properties are relatively uniform across the sample, the optimal imaging angle for each angular region lies exactly between the illumination and detection objective. In reality, the optimal imaging angle for each angular region does lie close to the midpoint between the illumination and detection angle, with a small variation. However, if we measure the optical accessibility for each angular region as

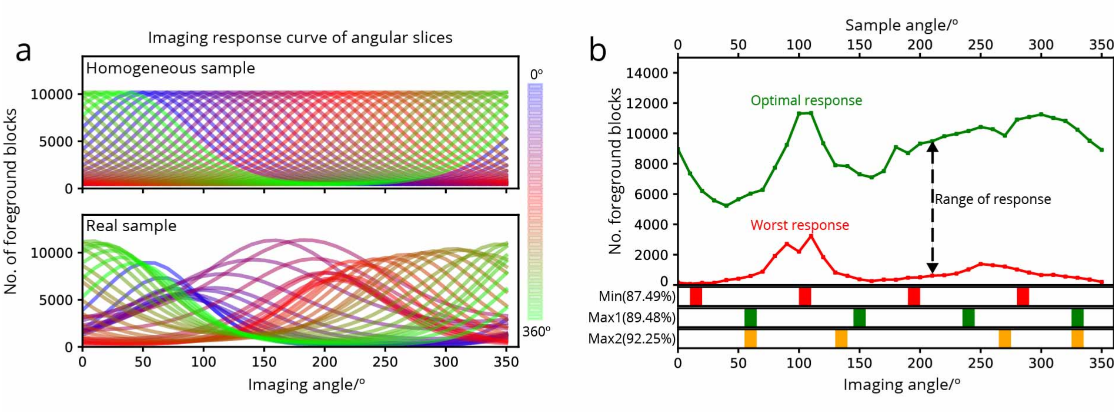


FIGURE 3.6 Evaluation of sample coverage profile. (a) Image response curve of angular slices comparison between a completely homogeneous sample and a real zebrafish embryo. Color codes the angular slice. (b) Top: Variation in imaging response for different angular region measured in the number of foreground blocks. Bottom: Comparison of sample coverage of 4 view multi-angle imaging strategies. Min: Minimum coverage given by 4 equally spaced views. Max1: Maximum coverage given by 4 equally spaced views. Max2: Maximum coverage with flexible view spacings.

the full-width half maximum of each fitted von Mises distribution, we can observe a considerable variance (**Fig** 3.7). In conclusion, the sample response profile can be used to evaluate both the fluorophore distribution and optical properties.

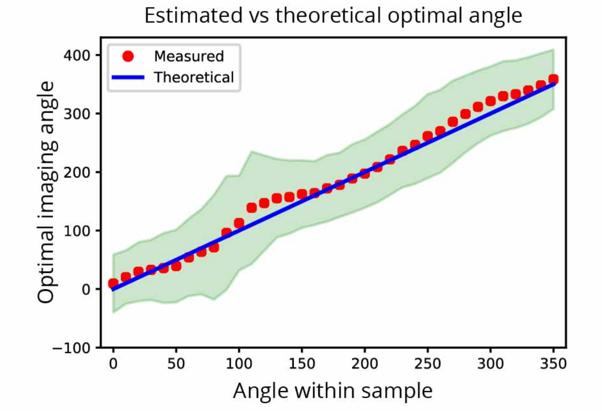


FIGURE 3.7 Optimal imaging angle for different angular regions of a zebrafish embryo. Shaded region represents the full-width half maximum of the fitted von Mises distribution.

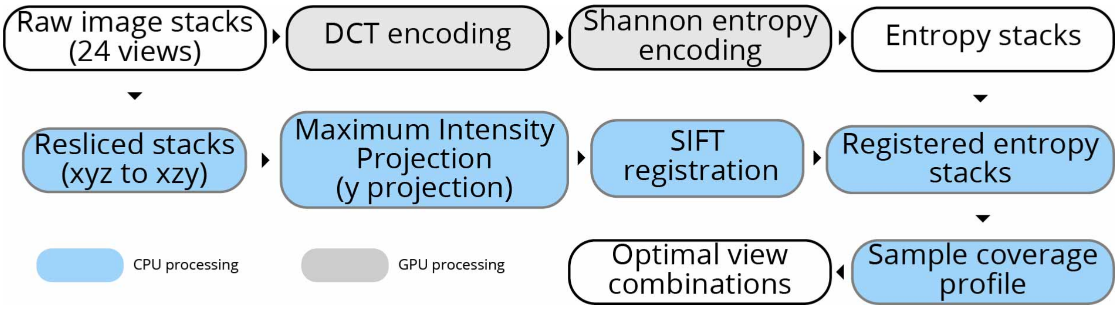


FIGURE 3.8 Smart rotation workflow.

3.2 SMART ROTATION WORKFLOW

3.2.1 ON-THE-FLY SAMPLE COVERAGE ESTIMATION

In addition to evaluating the sample's optical properties, the sample response profile can also be used to predict the overall sample coverage of different sets of views. I optimized the sample coverage estimation and sample imaging response profile generation such that the image analysis can be performed efficiently. I then built a custom microscope control workflow such that the image data can be processed as soon as they are generated and the predicted optimal view combination can be communicated to the microscope control software for reconfiguration (**Fig** 3.8).

To evaluate the imaging coverage accurately, the sample is imaged from N equally spaced angles. The number of angles N needed for evaluation is sample shape and labeling dependent. $N = 24$ angles were generally sufficient to generate a full sample response profile; using fewer angles gave less accurate registration between views, distorting the final coverage estimation. Image stacks are first processed on an Nvidia GPU to perform DCT encoding and Shannon entropy calculation (Royer et al., 2016; Shannon, 1948). I found CPU processing to be too slow for on-the-fly analysis, taking way more than 1 minute to process a single stack of images. The image stacks from different views are then registered for direct comparison. Full 3D registration between views is also too slow for on-the-fly applications. As there is only one rotational degree of freedom, I chose to summarize the image in the direction along the rotational axis. The maximum intensity projections of the raw image stacks along the rotational axis are registered with each other using a SIFT-based registration method (Saalfeld et al., 2010). The transformation is then applied to the minimum intensity projections of entropy image. The entropy image is used to estimate the number of foreground blocks based on a pre-determined background level. Automatic thresholding on the encoded image can also be performed, but I found that the entropy measure of information content is not absolute. A slight change in imaging conditions, including noise level and pixel exposure time, can drastically alter the entropy profile of the encoded image. The noise level on an sCMOS camera is very well controlled and consistent. A pre-determined background level can be approximated by the entropy of a blank image, which gives a much more consistent result. Foreground block counts are then summarized by angles with a bin size of 10 degrees to generate the angular information content profile. Similar to the number of views used, the bin size can be chosen based on the feature size of the sample. In this study, where a $10\times$ objective is used, a bin size of 10 degrees is sufficient to generate a full sample response profile. The information content profile is then fitted to a von Mises distribution (see **equation** 3.2) and both optical accessibility κ, and fluorophore distribution A, are calculated for each angular region. The performance for each view combination can be estimated as the average coverage percentage of all angular regions versus their estimated optimal coverage. Given the number of views to be imaged for each time point, the combination with the highest average coverage is used for imaging. The combined coverage of views is estimated in an additive manner.

It would also be possible to estimate the optimal view combination iteratively by taking only one view and then adding complimentary views with the most potential information gain. With sufficient prior knowledge of similar samples, the iterative estimation can perform comparably. Here I use 24 equally spaced angles during the estimation step to ensure applicability so that a complete information content map can be obtained for any sample type.

Although SPIM is amongst the fastest imaging modalities to generate full 3D image stacks, generating 24 views with ca. 500 frames each takes more than 2 minutes. Analyzing the 24 views to generate a complete sample coverage profile takes around 5 mins. The entire evaluation step takes roughly 6 mins if acquisition and evaluation are performed asynchronously rather than sequentially. In many *in toto* time-lapse imaging applications, such as zebrafish development studies, image stacks need to be acquired every minute or even faster. Therefore, it is not always feasible to perform a full 24-view evaluation for each time point during a time-lapse. Constantly image with 24 views also exposes the sample to higher photodamage, risking sample health. Instead of performing the full 24-angle evaluation at every time point, a full evaluation is performed at much longer interval and updates the set of views much less frequently.

In cases where subtle optical changes occur over a smaller time interval than the full 24-view evaluation step time interval, an update step can be used. During the update steps, the same analysis is performed on the last acquired views, and the overall fluorophore and optical accessibility map can be updated by substituting the corresponding views with the latest views. Since only the newly acquired views need to be processed, updates take a lot less time and can be run at every time point. In my implementation, the data acquisition runs independently to the image analysis workflow such that data acquisition is not affected in case analysis takes longer than expected. Before the start of each time point, the data acquisition computer communicates with the separate image analysis computer to query for the latest optimal set of views estimated by the workflow.

It is worth noting that down-sampling the image stack can significantly speed up the analysis as the DCT encoding time scales linearly with the number of voxels being processed. Too much down-sampling, however, can result in over-interpolation of the sample coverage profile. A compromise needs to be made to ensure the processing speed can keep up with the data acquisition rate.

The workflow utilizes the integrated data pipeline described in **Chapter** 2: Image data are streamed to a centralized file server, and all processing is performed on a powerful Linux workstation that is connected to the file server via 40 Gbps fiber optics connection.

3.2.1.1 Results

To verify that the smart rotation workflow improves the sample coverage in light sheet imaging, I imaged a 48-hour old zebrafish embryo (*Tg(h2afva:h2afva-mCherry)*). The sample was imaged with $N = 24$ equally spaced angles for the evaluation step which allowed me to compare the sample coverage performance between the blind multiview approach that uses equal spacing with a subset of views $n < N$ and the smart rotation workflow with the same number of angles n.

To directly compare the sample coverage between the two workflows, multi-view datasets were fused using multi-view reconstruction software (Preibisch et al., 2010), which utilizes a bead-based registration. I then performed information content estimation on the fused image stacks. Comparing the sample coverage between fused images from both workflows, the smart rotation workflow was able to consistently outperform the blind approach overall measured in relative coverage percentage (**Fig** 3.9).

The overall sample coverage increased non-linearly as the number of views n increased (**Fig** 3.10 a). There was a diminishing return in additional sample coverage as the number of views n increased. My smart rotation workflow converges to optimum

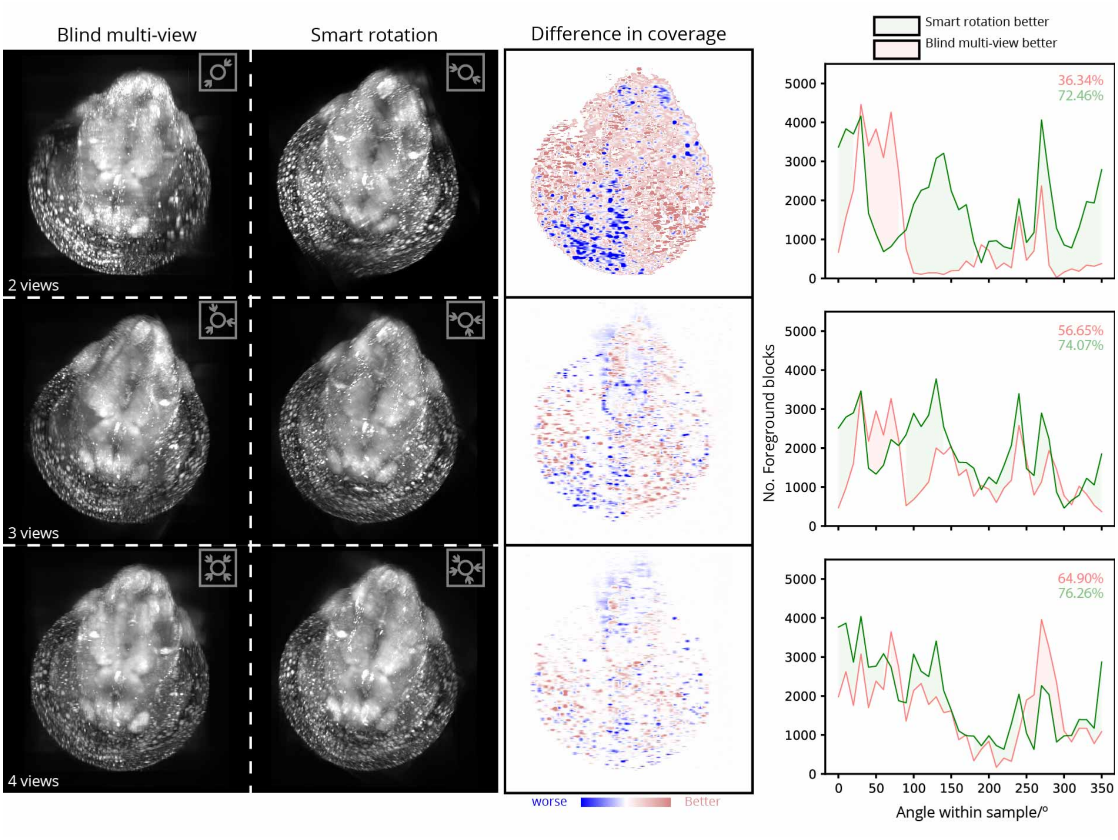


FIGURE 3.9 Performance comparison between blind multi-view and smart multi-view workflow. comparison between the fused images from a blind multi-view approach and our smart rotation workflow. The difference image is calculated as the difference in information content between the smart rotation generated image and the blind multi-view generated image. Number denotes relative image coverage.

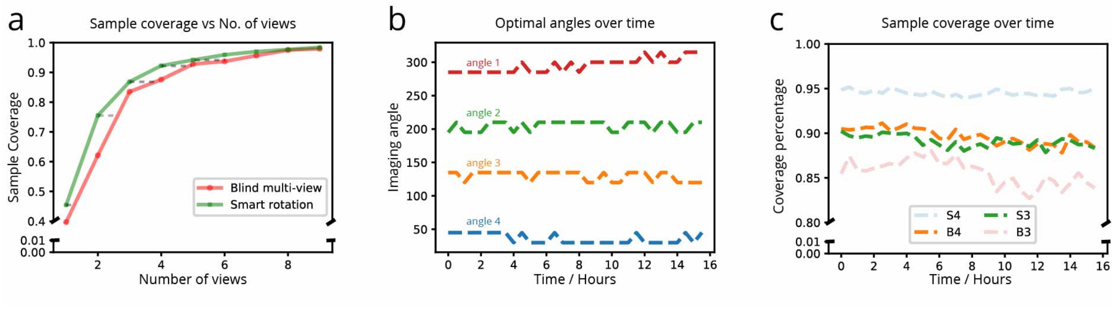


FIGURE 3.10 Evaluation of the performance of the smart rotation workflow over time.. (a) Sample coverage measured as a percentage of the maximum possible against the number of views used. (b) Angles selected by the smart rotation evaluation during 16 hours time-lapse of a zebrafish embryo (*Tg(kdrl: GFP)*). (c) Sample coverage percentage comparison between the smart rotation workflows and blind multi-view workflow over the 16 hours time-lapse. S denotes smart rotation workflow, and B denotes blind multi-view workflow, and the number denotes and the number of views used.

faster than the blind approach. In typical SPIM zebrafish experiments, 2 to 4 views are usually used for time-lapse imaging.

With my workflow, a similar or better sample coverage can be achieved compared to the blind multi-view approach with one less angle used (**Fig** 3.10 c). This directly translates to a reduction in the data volume and phototoxicity in the sample.

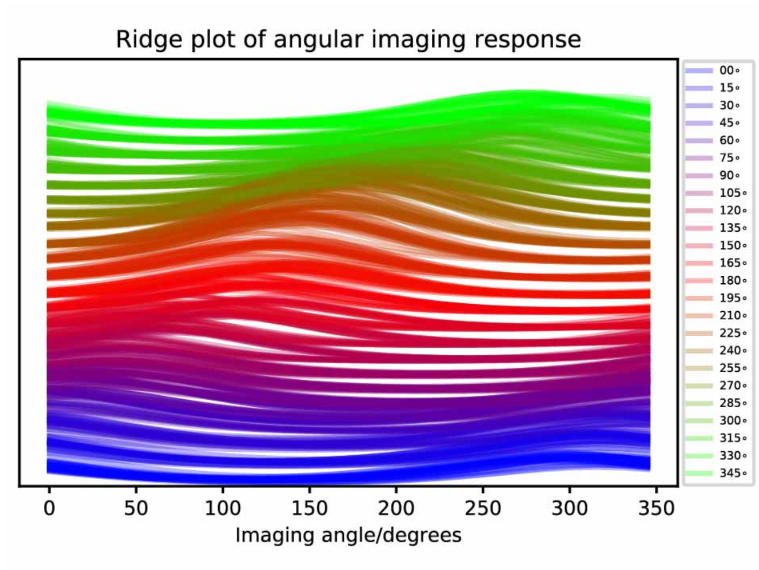


FIGURE 3.11 Ridge plot of estimated sample profile.

To verify that the smart rotation workflow can maintain the sample coverage over a time-lapse experiment better than the blind workflow, I imaged a zebrafish embryo with vascular labeling (*Tg(kdrl: GFP)*) from 48 hpf for 16 hours with 4 views. Evaluation steps were performed every 30 minutes. The 4 angles selected by the smart rotation workflow varied throughout time. The changes in each individual angle remained within 15 degrees (**Fig** 3.10 b). The angular image response evolved smoothly (**Fig** 3.11). More importantly, during a time-lapse experiment, the sample coverage in a blind multi-view workflow gradually decreases due to the sample's change in optical properties and fluorophore distribution. The sample coverage remained stable with the smart rotation workflow.

3.3 Conclusions

3.3.1 Discussion and outlook

I demonstrated the importance of smart multi-view imaging to improving sample coverage in fluorescence microscopy. The widely used blind multi-view imaging workflow did not yield optimal results (for a given number of views) and could have resulted in unnecessary phototoxicity. I formulated a method to evaluate the sample coverage in a multi-view experiment and used this metric to quantify the sample coverage differences between angular views in a living sample with inhomogeneous optical properties and spatial fluorophore distribution. After verifying the metric's performance with data generated from a real multi-view light sheet microscopy dataset, I built the coverage measurement method into a new smart multi-view workflow where optimal view combinations are estimated on-the-fly. I demonstrated that optimal imaging view combinations can be selected during acquisition using our smart rotation workflow. My workflow not only improves the overall quality of the images captured but also increases the amount of useful information in the data saved while reducing the phototoxicity.

Summarizing the information content distribution in 3D into angular slices is a significant simplification in the analysis workflow. It is possible to estimate the global optimum in 3D at pixel resolution since the angular imaging response for all regions is estimated. However, doing so would require a much more stringent 3D registration step, which typically cannot be performed during acquisition. I also tested running the workflow on images down-sampled in z, and I saw a minimal difference in the estimation of the information content distribution. The run-time of the analysis pipeline scales linearly with the number of z-frames, and therefore, the initial evaluation step can be accelerated by taking images at a lower z-sampling rate, reducing the total data volume. Lowering the z sampling also increases the speed of acquisition during the evaluation step. The amount of tolerable down-sampling depends on the intrinsic fluorophore distribution and therefore would need to be tuned for each sample type.

The analysis framework can be further accelerated if the data generated are analyzed in memory before saving to disk to eliminate file writing and reading time. Performing operations on image volumes in memory before saving to disk is usually only possible in home-built solutions. Therefore, I separated the data acquisition part of the workflow and the data analysis section such that users can write an interface layer to incorporate the workflow into their microscopes. In cases where on-the-fly image analysis is not possible due to hardware limitations, the workflow's evaluation step can still be used as standalone software to give a better estimation as to what stationary angles should be used in a time-lapse experiment. Currently, the smart rotation workflow is optimized for *in toto* imaging applications where information-rich voxels can be anywhere within the three-dimensional field of view. In these applications, different views are expected to cover a different area of the sample. However, it is also possible to give different weights to different areas of interest to optimize the image quality of a specific sub-region. I believe that the information content map itself is an effective representation of the sample that captures both the fluorophore distribution and optical properties. Therefore the map has the potential to be used as a template to estimate the optimal imaging condition for new experiments where similar samples have been imaged before.

To fully utilize the power of multi-view imaging, post-acquisition data fusion is required. In my experience, the image entropy changes significantly after image fusion. This can be attributed to the drastic change in noise characteristic. Besides, the widely used weighted average method of image fusion can decrease the overall image quality compared to the individual views: The well-imaged part of one view may be corrupted by another view with poor image quality in this region. In the future, we hope to extend the information content map to estimate how a specific view would affect the eventual fused result and adjust the weighting accordingly. Adaptive cropping of the data may also yield better fusion results, and will be more robust than manual cropping, which is sometimes done.

There is one other benefit of multi-view imaging: Multi-view imaging can rotate the system point spread function (PSF) relative to the sample. As a standard 2-lens SPIM has different axial and lateral resolutions, imaging the sample with multiple PSF orientation enables multi-view deconvolution to improve the spatial resolution (Preibisch et al., 2014; Schmid and Huisken, 2015). The deconvolution result depends heavily on the correct estimation of the PSF. The smart rotation workflow can also help improve the deconvolution result by offering raw image stacks with better image quality.

3.3.2 INFORMATION CONVERSION EFFICIENCY IMPROVEMENT

The smart rotation workflow intuitively improves the information conversion efficiency by increasing the amount of information contained in the raw image stacks (increasing the numerator of **equation** 1.1). The improvement can be directly measured by comparing the information content maps between the blind workflow and the smart workflow. The relative improvement is often minor (**Fig** 3.10 c). However, since in the smart rotation workflow, it is possible to reduce the number of views needed to achieve the same eventual sample coverage, the raw data size can be significantly reduced. In the case shown above, the performance of 3 views in the smart rotation workflow is comparable to that of 4 views in the blind case. A 25 % increase in information conversion efficiency can be observed.

4

Integrated SPIM workflow for neuro-developmental imaging

> *"As technology advances, it reverses the characteristics of every situation again and again. The age of automation is going to be the age of 'do it yourself.'"*
> — *Marshall McLuhan*

In the previous two chapters, I described methods to improve the information conversion ratio (see **equation** 1.1) in order to address the data burden in SPIM. In this chapter I will demonstrate an implementation of the third method: direct raw image data to data product conversion pipeline to alleviate intermediate storage stress.

4.1 Introduction

A neuron is one of the most morphologically interesting cell types. Neurons develop from a simple spherical shape into a complex structure with elaborate axonal and dendritic projections. The astounding developmental progression relies on robustly regulated assembly and dynamics of the cellular cytoskeleton (Coles and Bradke, 2015; Flynn, 2013; Menon and Gupton, 2016; Pacheco and Gallo, 2016). Neurons can extend their axonal projections many times the length of their cell body, in some cases up to $10,000$ times the length of the initial cell body. Precise regulation of cargo trafficking and axonal transport is crucial for neuron development, function, and maintenance (Nirschl et al., 2017). Cargos are delivered to their target locations via a diverse cohort of kinesins and their cargo linking adaptor proteins.

It has previously been shown that the kinesin-1 adapter, Calsyntenin-1 (*clstn1*), has important roles in cargo transport, microtubule polarity, and axon branching in sensory neurons (Lee et al., 2017; Ponomareva et al., 2014). Axon branching is an important mechanism for axon growth and sensory field coverage. Clstn1 binds to kinesin light chain (*KLC*), a subunit of kinesin-1, which then becomes activated for cargo transport (Yip et al., 2016). However, there are still major unanswered questions such as: What is the mechanism of *clstn1* action and what are the other associated but unknown players in axon branch regulation? One of the candidate genes is *KLC* as it is directly connected to *clstn1* and is well understood to have a precise role in the kinesin motor machinery.

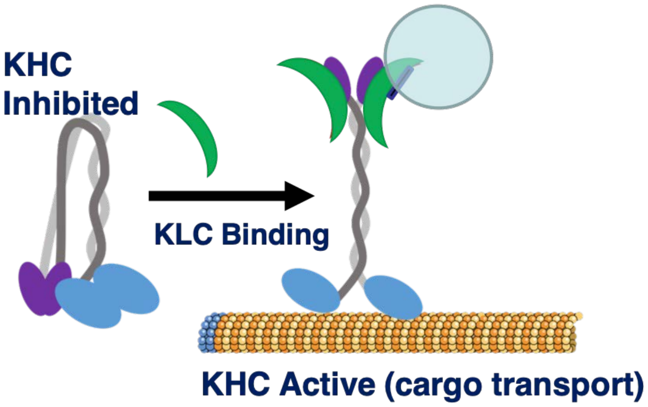


FIGURE 4.1 Illustration of Kinesin. Illustration by *Dr. Elizabeth M. Haynes*.

4.1.1 THE ROHON-BEARD CELL MODEL

To study the axonal branching pattern and how it is related to specific genes, we use the Rohon-Beard (RB) cell in zebrafish as a model.

The RB cell is a type of sensory neuron in early fish and amphibian development. They are found near the dorsal part of the spinal cord and have mechanoreceptive properties. During the first 48 hours of development, RB cells grow out elaborate axonal projections close to the surface of the embryo.

The RB cells are positioned in a linear pattern along the anterior-posterior axis within the spinal cord. The cell bodies' initial neurites fasciculate with neighbors to form a long extended main fascicle or central axon before around 16 hpf. Then each cell sends out peripheral axonal projections close to the surface in a dorsal-ventral orientation(**Fig** 4.2). After around 32 to 33 hpf, the axonal projections would cover most of the body of the zebrafish embryo (**Fig** 4.3).

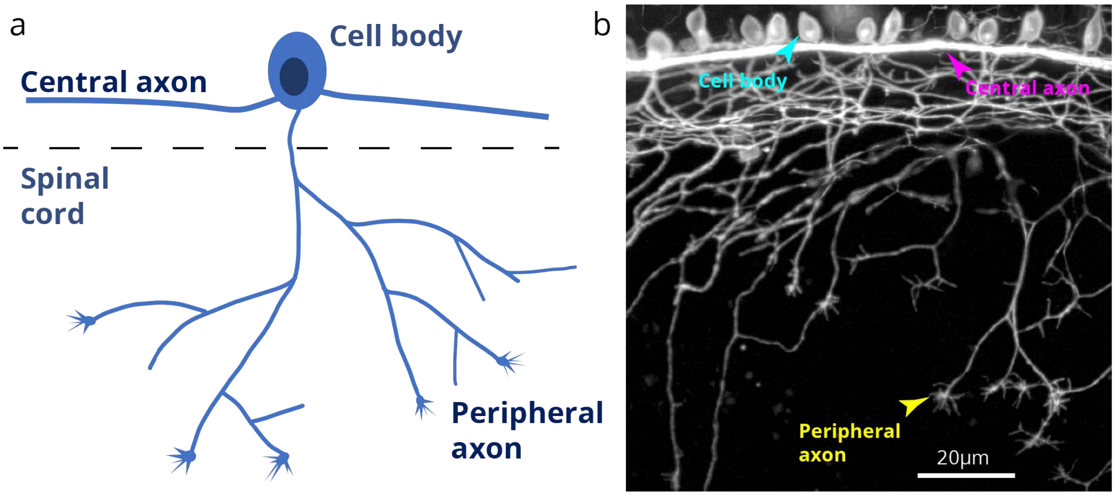


FIGURE 4.2 (a) Illustration of the anatomy of a Rohon-Beard cell in zebrafish embryo by *Dr. Elizabeth M. Haynes*.(b) Image of RB cells in a zebrafish embryo (*ngn1:GFPcaax*) at 20 hpf.

RB cells only exist in early embryonic stages. They typically disappear through programmed cell death after 2 to 4 weeks in zebrafish and are replaced by dorsal root ganglion (DRG) of spinal nerves (Kanungo et al., 2008; Reyes et al., 2004). The behavior of the RBs is analogous to other sensory neurons and the mechanism of axon branching reflects other neuron populations. Therefore it is a good model to study general axon behavior.

Since the axonal projections of RB cells are very superficial, they are optically very accessible, making them ideal for imaging experiments (Andersen et al., 2011).

4.1.2 LIVE IMAGING OF RB FOR PHENOTYPIC QUANTIFICATION

RB cells have primarily been imaged with confocal microscopy in previous studies (Lee et al., 2017). However, the imaging speed of point-scanning confocal microscopes is very limited, resulting in low experimental throughput. In order to elucidate the functions of specific genes in RB axon branching, we use SPIM to perform phenotypic screening. With the help of our collaborator (*Dr. Elizabeth M. Haynes* and *Professor Mary Halloran*) at the University of Wisconsin-Madison, we gained access to zebrafish lines that are genetically perturbed using the CRISPR/Cas system. By quantifying the phenotypic changes associated with the specific perturbation, we can further elucidate the mechanisms governing axonal branching.

There has been no study that performs 3D fluorescence microscopy based phenotypic

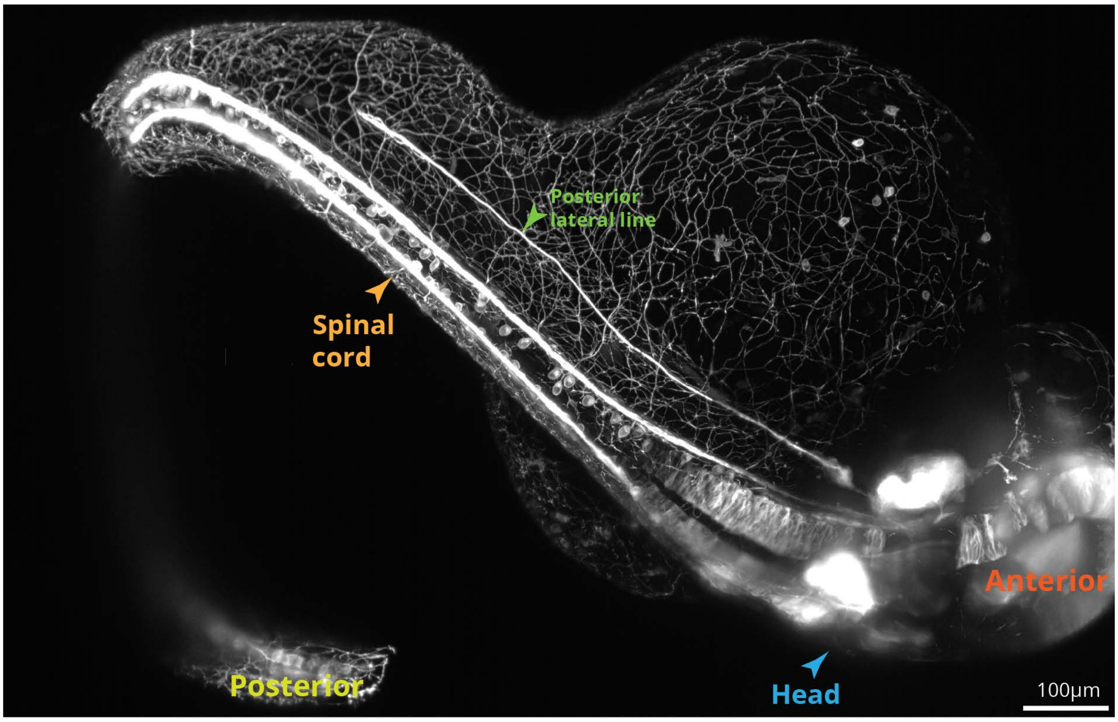


Figure 4.3 Zebrafish embryo (*ngn1:GFPcaax*) imaged at 28 hpf. The axon projection covers a significant portion of the sample body. The ngn1 promoter labels, in addition to the RB cells, the mid and hindbrain, neuroepithelial cells, neural crest cells, trigeminal and posterior lateral line.

screening of zebrafish embryos with reasonable throughput. There have been efforts to build screening compatible sample holder for multiple zebrafish embryos for confocal live imaging (Megason, 2009). However, the speed of a conventional confocal microscope limits the potential throughput of the screening procedure. Even if a faster confocal microscopy variant is used (such as spinning disk confocal microscopy), the high photo-toxicity limits the potential to generate time-lapse to quantify the temporal dynamics of axon development. In comparison, SPIM is much faster and less phototoxic, ideal for higher throughput 3D imaging.

We aim to create a new integrated SPIM-based screening system for early zebrafish embryo. The system consists of Both a newly designed mid to high-throughput 3D imaging SPIM and an on-the-fly quantification pipeline that can correctly identify phenotypic differences of samples with different genotypes.

4.2 Custom integrated SPIM workflow

Before constructing a custom high-throughput imaging system, I started with a pilot study to image RB with relatively low throughput using a standard SPIM system. I used the same camera and similar imaging objectives in my SPIM to the eventual system to ensure that the data generated are comparable. The goal of the pilot study is to establish a phenotype quantification pipeline that is capable of correctly separate samples with known axon projection branching defects to wild types through image analysis.

4.2.1 Challenges

There are many challenges involved in designing a new microscopy experiment workflow. First of all, an optimal sample preparation protocol is critical for maintaining sample health. The amount of physical constraint the imaging environment imposes upon the sample can have huge impacts on the imaging result.

Data handling also needs to be taken into considerations. SPIM is capable of generating *3D+t* images at high framerate. As previously mentioned in **Chapter** 2, data throughput is directly related to the experiment throughput and therefore has to be taken into consideration.

Although automated image workflow to handle SPIM image analysis has been published, they are often very computationally demanding and cannot be performed during image acquisition (Daetwyler et al., 2019; Schmied et al., 2015). In our study, multiple samples need to be imaged in a single experiment, but not all image stacks carry the same amount of information. Phenotypically interesting samples need more attention than those that are identical to wild type (WT) samples. The result generated by immediate image quantification can be used to determine if it is meaningful to continue image the same sample or move on with other ones.

In this section, I detail the complete imaging pipeline used in the pilot study from sample embedding to on-the-fly quantification.

4.2.2 Optimal sample embedding in a standard SPIM

Since our imaging window is around 16 hpf to 40 hpf during yolk sac and body extension, the gel must not be too stiff to counteract the growth. Our lab has previously

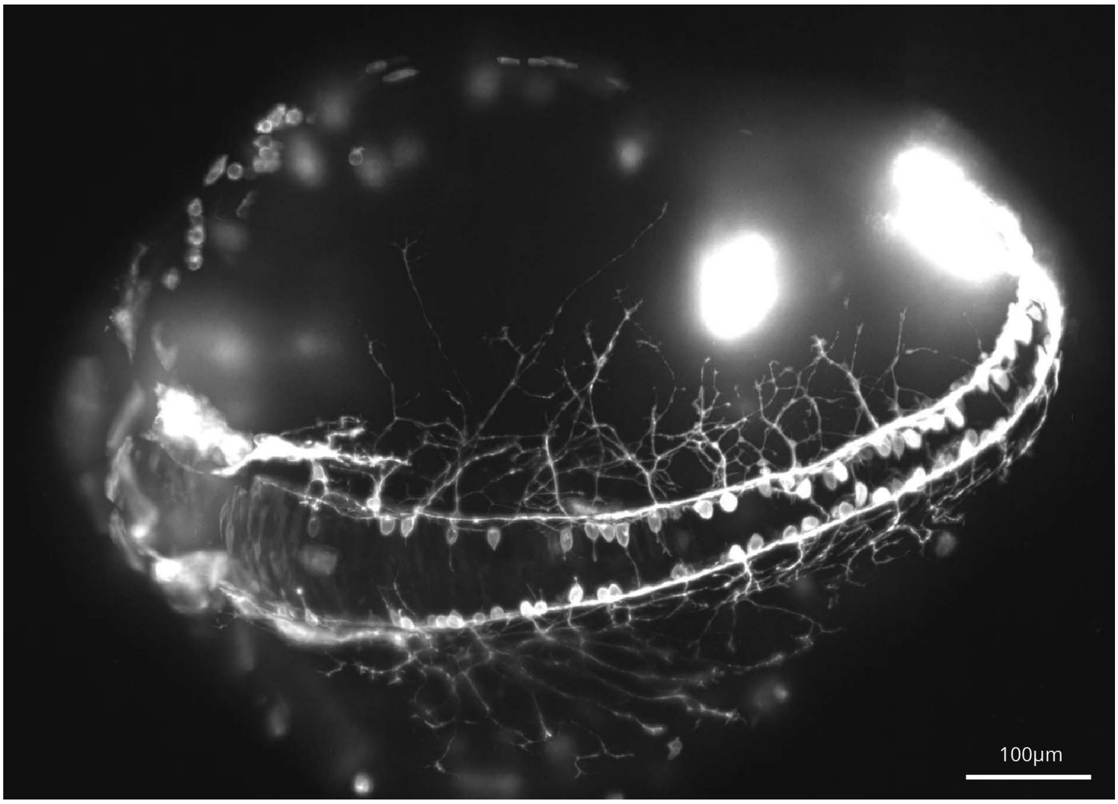


Figure 4.4 A zebrafish embryo (*ngn1:GFPcaax*) is embedded with 0.8% low melting point agarose in an FEP tube from 16hpf, imaged at 20 hpf, shows a balled-up embryo without natural tail extension.

characterized the optimal sample embedding parameters extensively for zebrafish in a standard SPIM system (Kaufmann et al., 2012). Typically, the sample is held vertically within a fluorinated ethylene propylene (FEP) tube that is refractive index matched with water. For SPIM time-lapse volumetric imaging, maintaining the sample to stay within the field of view is critical. Therefore, the sample is usually embedded in the tube with low concentration agarose to provide support. However, using agarose to hold the sample in place also puts external forces on the embryo that could potentially damage the sample. Hence, depending on the application, a compromise between agarose rigidity and sample growth is needed. For example, in zebrafish functional cardiac imaging, the sample is only imaged for a short duration at high speed: a concentration of 0.5% is typically used. For our RB development study, since the axonal projections are very superficial, it is crucial that they are not subject to external pressure to ensure healthy growth. However, it is also important to ensure the imaging view captures as many axon projections as possible. We initially tested embedding the sample with up to 0.8% low melting point agarose (*Sigma A9414*). A noticeable developmental defect can be observed when the gel impacts the tail extension (**Fig** 4.4).

Therefore, we decided to embed the sample in FEP tubing with only the E3 media. The sample is instead only held up by a 1.5 % agarose plug as previously described (Kaufmann et al., 2012; Weber et al., 2014).

As the sample is not constrained by gel, it can freely develop within the FEP tube. As a result, the eventual sample orientation relative to the detection orientation becomes arbitrary. Since the axon branching is very superficial, imaging the lateral view of the sample should capture most axon branches. Hence, it is necessary to readjust the imaging view through post-processing such that an ideal lateral view can be visualized.

4.2.3 Custom image analysis software for sample phenotyping

4.2.3.1 Motivation

In order to perform image analysis based phenotyping, we need to create a robust and efficient processing software. Neuron image analysis is a very well studied subject. Neuron segmentation and reconstruction from electron microscopy or fluorescence microscopy images have always been challenging tasks. There even exist academic competitions to boost the accuracy and speed of neuron reconstruction (Liu, 2011). However, the majority of the best performing neuro-tracing methods are only semi-automatic and requires a significant amount of human annotation. The users usually either define key landmarks of each neuron by hand or manually proof-read the often inaccurate automated tracing result (Peng et al., 2011). The process is highly time-consuming and labor intensive. For our screening application, it is less important to trace each neuron accurately than to find significantly distinguishing features between sample types. Since we are interested in the axonal branching pattern of RB cells, we hypothesized that the branching orientation relative to the main fascicle is different between wildtype embryos and embryos perturbed either genetically or pharmacologically.

4.2.3.2 Sample feature assisted image unwrapping

In a standard confocal experiment, zebrafish embryo is embedded on a bed of agarose and can be manipulated with a thin pipette such that a perfect lateral view of the sample can be recorded. However, since the sample is freely embedded in a vertical FEP tube in SPIM, it is difficult to align the lateral orientation of the fish with the

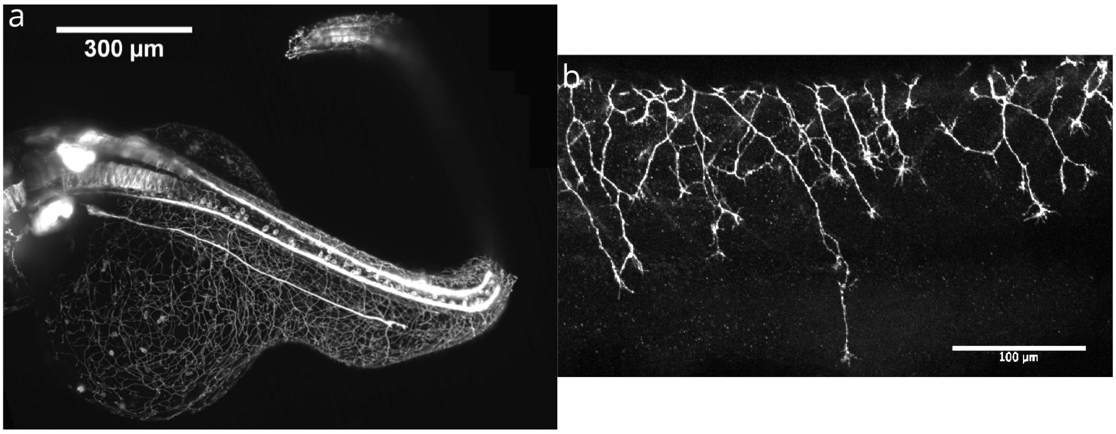


FIGURE 4.5 (a) Sample imaged at 24 hpf in a standard SPIM. The sample is not embedded with any agarose. (b) Sample imaged at 24 hpf in a confocal microscope. The sample is embedded flat in a petri dish and re-oriented so that the lateral view faces the imaging objective.

image plane. As a result, the sample often appears curled (**Fig** 4.5) or tilted. To be able to quantify the orientation of axon branching relative to the direction of the main fascicle, it is essential first to adjust the axis orientation in the 3D image stack.

4.2.3.2.1 Finding body axis To convert an image stack of a fish in any orientation to a lateral view, I need to straighten the sample in image space. Christensen et al., 2015 has performed similar tasks to straighten a *C. elegans* embryos. They selectively expressed fluorescence proteins in certain landmark cells to extract the body axis. In our case, we can also exploit features specific to RB cells.

The main fascicle is exactly along the spinal cord of the sample and is therefore an excellent estimate of the body axis of the sample. We use *ngn1:GFPcaax* zebrafish which has fluorescence in the membrane of RB cells. Conveniently, the main fascicle has the highest fluorescence expression level and therefore is the brightest part of the image. No complex segmentation is needed other than automatic thresholding to extract the brightest point in each image slice. The spatial coordinates are then fitted with a smoothed curve as an estimate of the body axis (**Fig** 4.6).

4.2.3.2.2 Sample straightening After the main body axis has been extracted, a reference frame is created along the axis. There are many ways to create a canonical reference frame along a given curve. Previously, the Frenet-Serret frame has been used to create a reference frame along the mid-line of a zebrafish embryo heart (Weber et al., 2017). The heart can then be unwrapped by reslicing the images along the normal direction

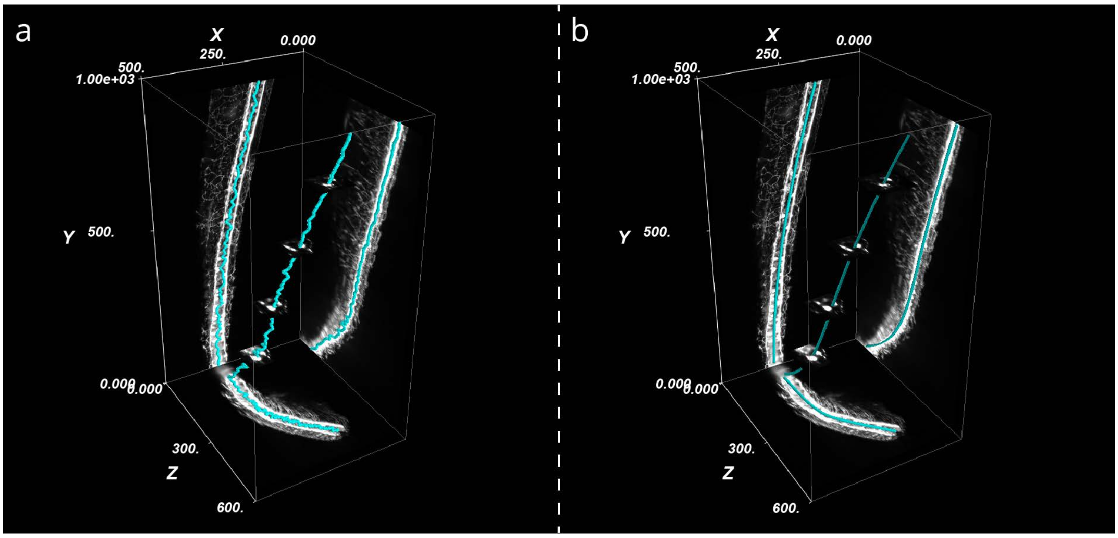


FIGURE 4.6 (a) Brightest part of each image plane align with the main fascicle and the body axis direction. (b) Connecting and smoothing the curve gives an estimate of the body axis

of the reference frame. However, the Frenet-Serret frame has a main disadvantage. Consider the formulation for the Frenet-Serret frame:

$$\frac{d\vec{T}}{ds} = \kappa\vec{N} \tag{4.1}$$

$$\frac{d\vec{N}}{ds} = -\kappa\vec{T} + \tau\vec{B} \tag{4.2}$$

$$\frac{d\vec{B}}{ds} = -\tau\vec{N} \tag{4.3}$$

where $\vec{T}$ is the tangential vector, s is the unit arclength defined along the curve. $\vec{N}$ is the normal vector. $\vec{B}$ is the binomial vector. κ is the curvature of the axis,τ is the torsion of the axis. The direction of each vector is estimated at each control point.
Since the normal direction, $\vec{N}$ can be taken arbitrarily when the tangent direction $\vec{T}$ has a vanishing second derivative, there is no guarantee of continuity between control points (Kreyszig and Pendl, 1975). As a result, between control points, random re-orientation of the normal vector can be observed.

Instead, we use an alternative method called parallel transport to estimate the reference frame along the axis (Hanson and Ma, 1995). In the parallel transport frame, the tangential vectors, $\vec{T}_I$ at all the control points, i, are calculated in the same way. Since the definition of the two other vectors can be chosen arbitrarily within the normal plane, we can transport the normal and binormal vectors along the curve to ensure continuity. After an initial normal vector is calculated as $\vec{N}_0$, which is perpendicular to

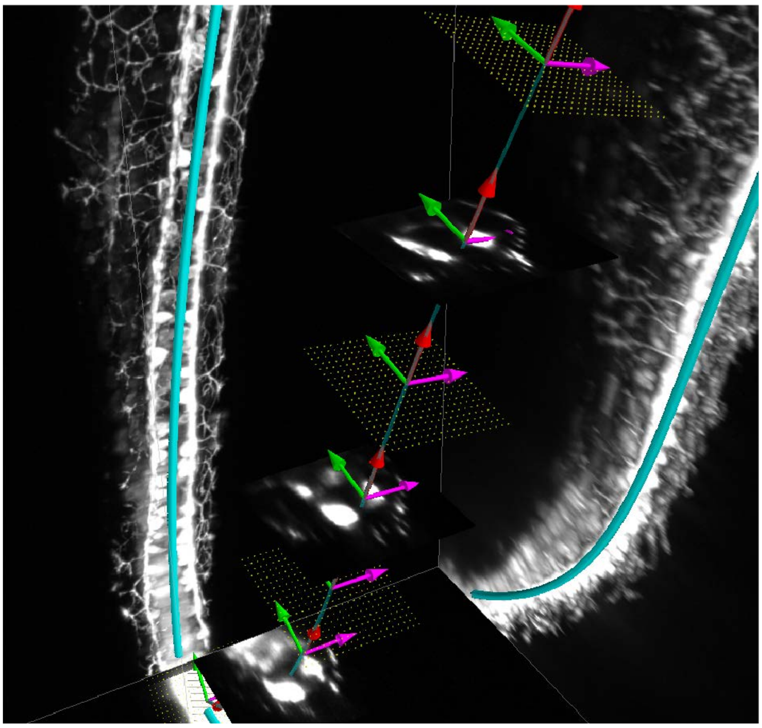


FIGURE 4.7 The image is resliced along the estimated body axis and reinterpolated into a straightened frame. Red arrows denote the tangential vector $\vec{T}$, green arrows denote the normal vector $\vec{N}$ and the cyan arrows denote the binormal vector $\vec{B}$.

$\vec{T}_0$, we can calculate the binomial vector as follows:

$$\vec{B} = \vec{T}_i \times \vec{T}_{i+1} \tag{4.4}$$

If $||B|| = 0$, $N_{i+1} = N_i$. Otherwise $N_{i+1} = R(\hat{B}_i, \theta)N_i$. Where $\hat{B}_i$ is the unit vector in the direction of B and θ is the angle between T_i and T_{i+1}.

The binormal vector always depends on the previous control point. Effectively, we reduce the degree of freedom of the axis orientation by 1 and ensure continuity in the axis orientation. After the reference frame is estimated, the images can be resliced along the body axis of the fish, therefore straightening the image (**Fig** 4.7,4.8).

4.2.3.2.3 Image unwrapping After the sample is straightened, the sample is ready to be analyzed in an ideal lateral view. However, since the RB cells are mainly superficial, the fluorescence signal only comes from the surface. Therefore, the majority of the voxels does not contain any useful information for analysis. Moreover, since the

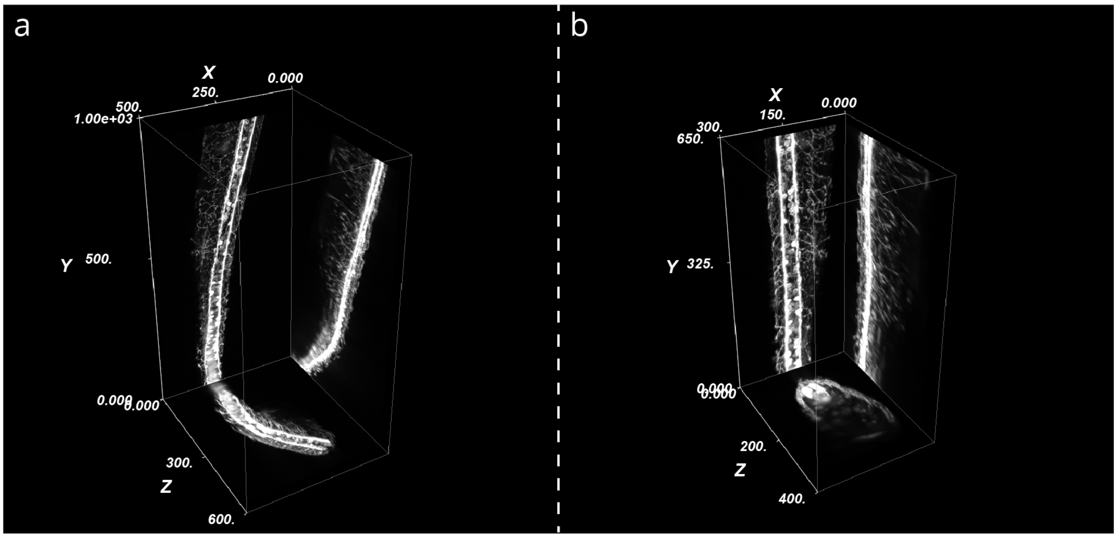


Figure 4.8 Image straightening along the fish body axis. (a) Original raw image where the sample's body axis is arbitrary relative to the image frame. (b) Straightened image with the anterior-posterior axis align with the axial direction of the image stack.

body of the zebrafish is not a flat surface, treating the lateral view of the sample as a flat surface is likely to introduce system errors when calculating the orientation of branching. To appropriately extract the surface where the axonal projections live on, I implemented a surface unwrapping software. The software outputs a 3D stack of the surface containing the axonal projections.

The first step of the algorithm is to estimate the surface of the embryo that contains the axon projections. I only consider the yolk extension segment of the zebrafish body as it is a good landmark for reproducible ROIs. The cross-sectional area of the fish body can be estimated as an ellipse, and therefore, the surface is a cylinder off elliptical cross-sections with varying diameters. I fitted an ellipse to each plane perpendicular to the body axis direction. The fitted ellipses are then combined and interpolated to form the surface of the fish body (**Fig** 4.9).

As the signal lies primarily on or near the surface, I can then project the data onto the surface and map it directly to a 2D plane. Similar work has been done to extract surfaces within a 3D image stack (Heemskerk and Streichan, 2015; Schmid et al., 2013). The term 'tissue cartography' is used to describe the method of reducing the dimensionality of the data without losing crucial details. Here the mapping between a cylindrical surface to a 2D plane is trivial as the two are topologically equivalent. Both distance and angle are preserved during the mapping. Since only a small proportion of the pixels on the surface of the embryo are used for unwrapping, the unwrapped

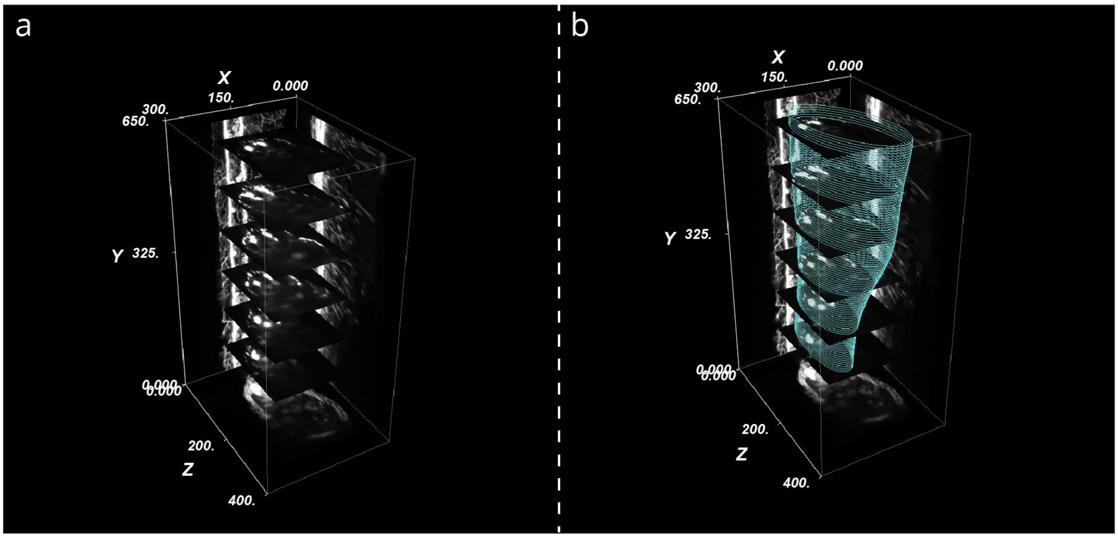


FIGURE 4.9 (a)The straightened image stack has all slices along the body axis aligned with the image axis. (b) An ellipse is fitted to each image slice, and together they form the surface of the sample.

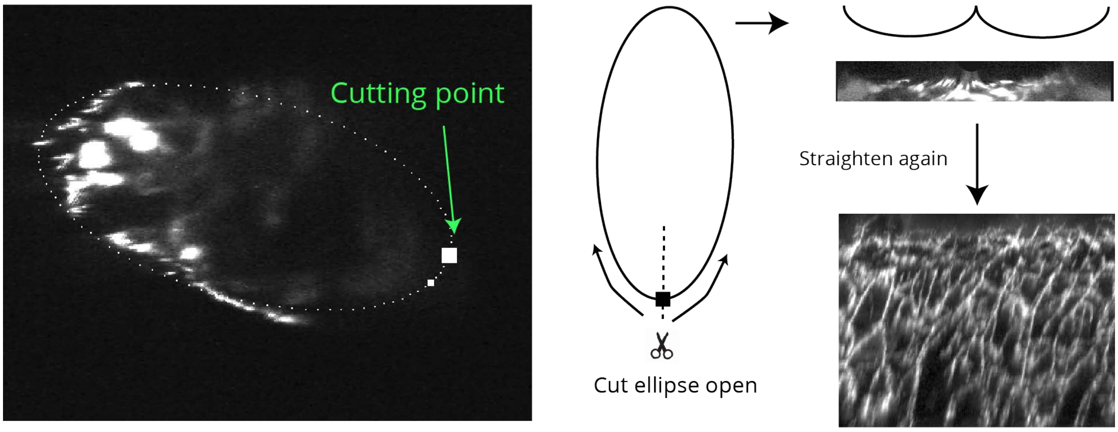


FIGURE 4.10 Illustration of the sample unwrapping method.

data is about 1% of the original full stack size.

4.2.3.2.4 Orientation analysis To quantify the orientation of the axonal branching, we used existing orientation analysis software (Püspöki et al., 2016; Tinevez, 2017). The method is pixel-based and gives a profile of the axon orientation relative to the central main fascicle. To clean up the image and enhance the axon features, I also apply a tubeness filter to the image. Orientation analysis is performed with *Fiji*.

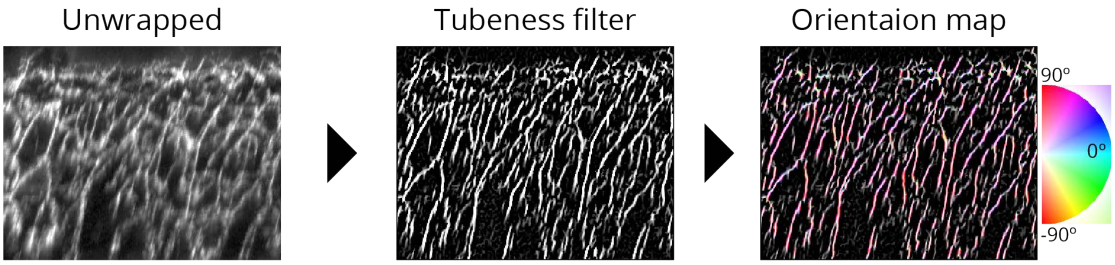


FIGURE 4.11 Orientation analysis is performed on the tube filtered unwrapped image.

4.2.4 RESULTS

To show that the image analysis is capable of identifying phenotypic differences, I needed a positive control that has a known phenotypes identified by traditional confocal microscopy. *CK666* is a molecule that inhibits the actin related protein (*Arp*) and is known to be important for axon branch formation (Hetrick et al., 2013; Ketschek et al., 2015). Samples treated with *CK666* show a preferential branching direction towards the posterior part of the embryo whereas WT embryos have more even distribution of branching direction (**Fig** 4.12).

It is common to first try image analysis on the maximum intensity projected (MIP) images as they are 2D and easier to manipulate. If orientation analysis is run on the MIP images, the arbitrary orientation of the body axis dominates the analysis result, and the drug-treated sample and WT are less distinguishable. When running the orientation analysis on the unwrapped image, the profile differences between the two different samples become more pronouced(**Fig** 4.13).

Additional verification is needed to confirm further that the orientation profile can be used to identify other known phenotypes as well. We notice that there are other features such as average branching angle and distances between branch points that appear to be different between WT and perturbed samples. I plan on testing image analysis methods such as 2D neuron tracing on the unwrapped image to see if more information can be extracted. In the future, we plan on using a wide variety of processing heuristics to generate a more complex profile for more precise phenotyping.

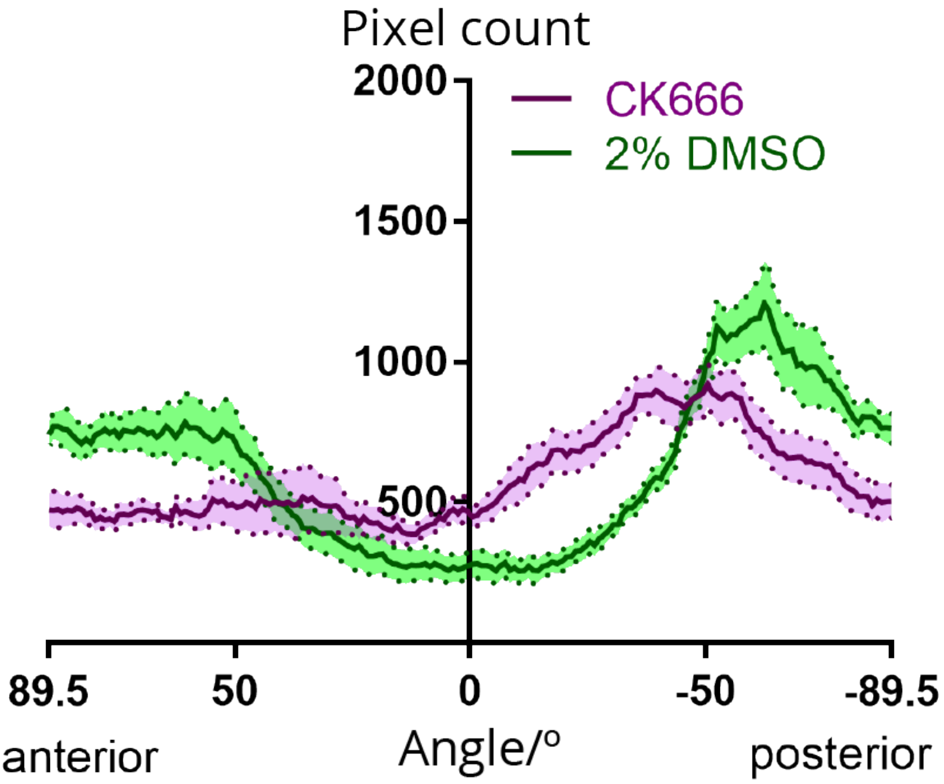


Figure 4.12 Orientation analysis of images taken with a confocal microscope in lateral view. Samples are either treated with *CK666* or 2 % *DMSO*.

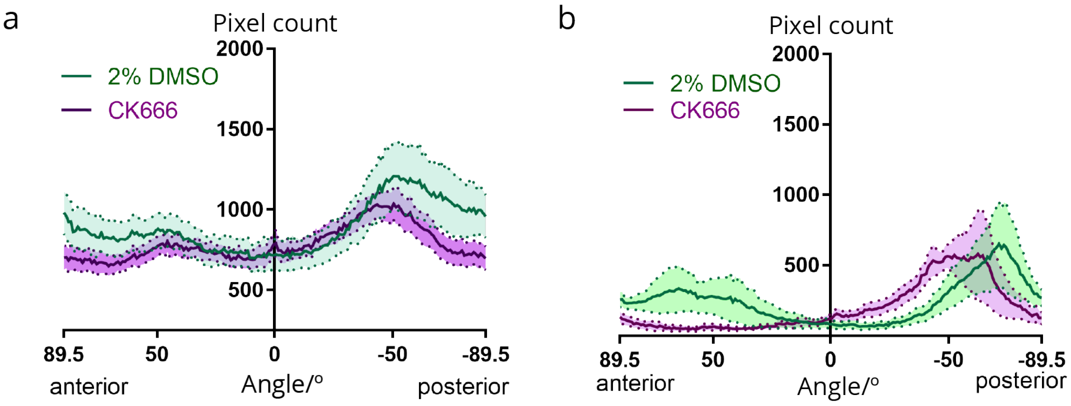


Figure 4.13 Orientation analysis of images taken with a SPIM in an arbitrary view.

4.3 MID-THROUGHPUT SPIM BASED SCREENING SYSTEM

After establishing the integrated SPIM workflow, I then set out to improve the throughput of our system. With our conventional vertical SPIM system, we can only embed one fish per FEP tube. To image multiple zebrafish embryos in a single screening experiment, each sample would have to be embedded into different tubes and image sequentially with manual sample exchange. It is therefore impossible to automatically screen fish at different time points.

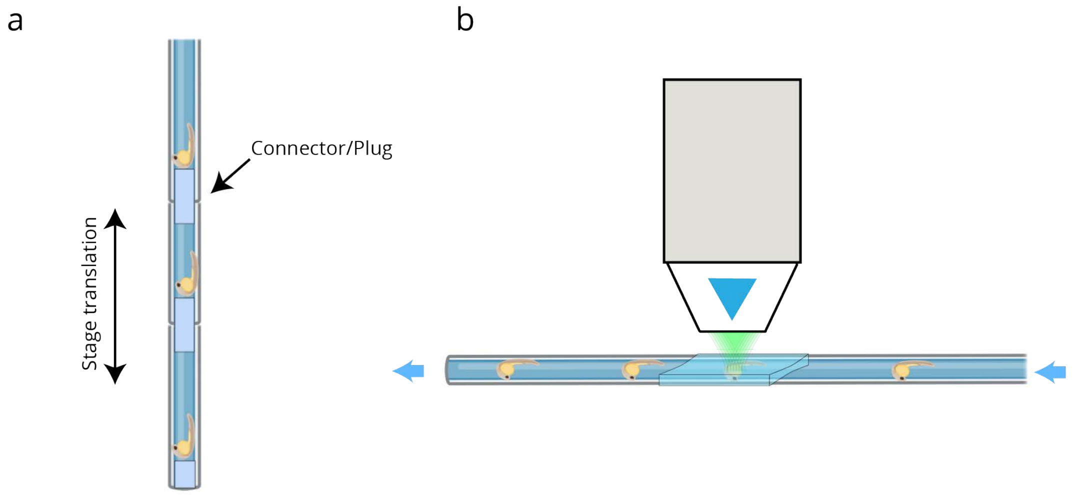


FIGURE 4.14 Alternative mounting for improved SPIM throughput. (a) Multi-sample mounting of up to 5 embryos in a single experiment by connecting multiple FEP tubes with an adapter. (b) High-throughput sample loading via fluidic control of sample positions in a single tube.

There have already been efforts to increase the throughput of a standard SPIM system. Daetwyler et al., 2019 use adapters to daisy-chain multiple FEP tubes together and can image five embryos in a single experiment. Shah, 2016 placed zebrafish embryos in a continuous FEP loop and used fluidic devices to control the position of each embryo. As a result, multiple embryos can be imaged sequentially.

Both approaches increase the throughput of a vertical SPIM system but also have several limitations. The first approach is not scalable to even higher throughput. As the number of tubes increase, the overall length of the assembly and the necessary travel distance increase. Also, an extended hanging tube has worse stability compared to a shorter tube, potentially affecting imaging resolution. *Shah et al.* also showed stacking multiple embryos close together without any connectors. However, this configuration is only applicable to early fish embryos in their chorion. In the second case, the sample cannot be oriented within the tube and sample positioning within the tube might be suboptimal.

4.3.1 Mid-throughput zebrafish mounting for V-SPIM

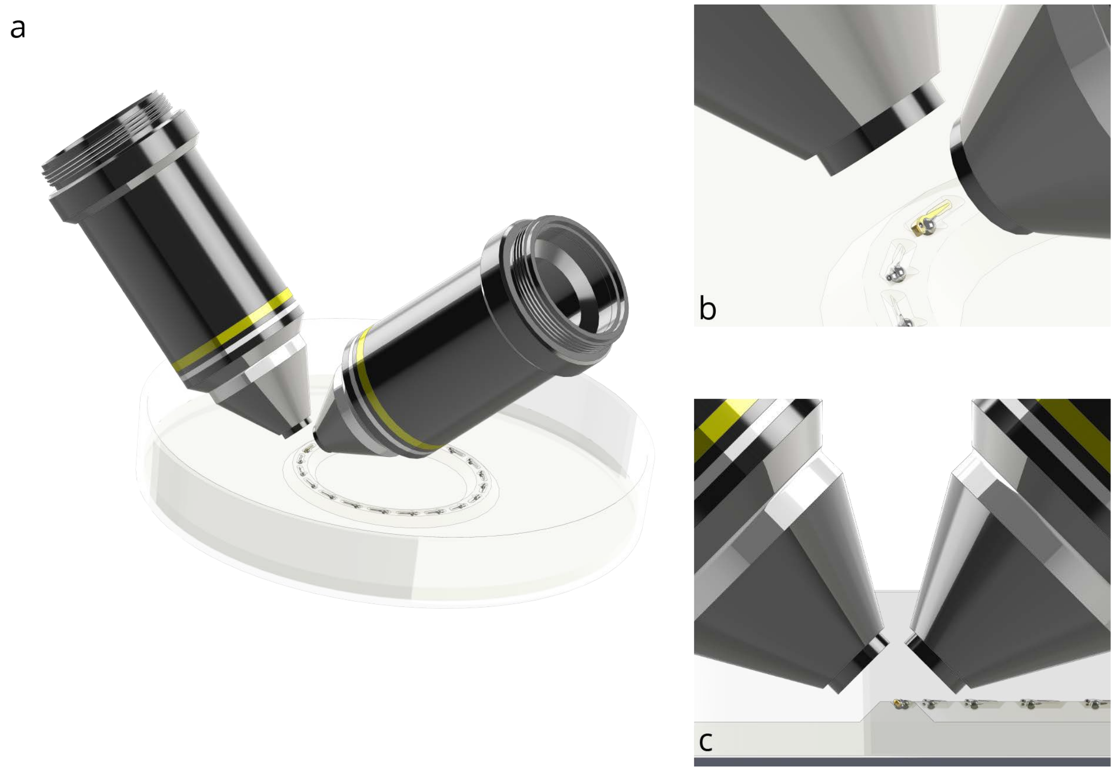


Figure 4.15 (a) Rendering of a mounting scheme where samples are mounted in a circle in a petri dish. (b) The illumination and detection objectives come from above. Only a small rotation is needed to move the FOV from one sample to the next. (c) The sample sits in a pocket in the agarose such that the lateral orientation faces the detection objective. Rendering generated by *Todd Bakken*.

To make it easier to mount multiple samples for a single experiment, we choose to build a SPIM that adopts the geometry of an iSPIM (Wu et al., 2011) or V-SPIM . The geometry enables a more conventional Petri dish based mounting for biologists. We developed a 3D printed mold to shape the agarose in the dish. The shaped agarose have specific pockets for each embryos to stay in. The embryos are arranged in a circular pattern. Similar to the *megamount* geometry described in Megason, 2009, our pattern enables efficient embedding and imaging of multiple samples.

Since SPIM utilizes two objectives for illumination and detection separately, the distance between neighboring samples on the grid needs to be large enough to avoid contact between the objectives and a neighboring sample. The longer the grid spacing is, the larger the distance the stage have to cover between samples. Therefore we designed a circular pattern plate which eliminates the sample access issue and is capable of holding 25 samples (**Fig** 4.15). A rotational stage is used to rotate the Petri dish in order to move from one sample to the next.

The circular mounting pattern can be imaged much faster than the other two methods. When using the long extended FEP tube method, the translation stage needs to move a significant distance (larger than 3mm) between samples. In the case of the fluidic control system, the speed of the flow is limited by the inertia caused by the fluid movement. In our case, high-speed rotational stage such as *PI U-651* can move at up to 720° per seconds which means the moving time between samples can be as little as 20 milliseconds (**Table** 4.1).

Methods	No. of samples	Translation time (sec)	Disadvantage
Daetwyler et al., 2019	up to 5	2-3	Instability
Shah, 2016	unlimited	2-3	Difficult preparation
Our method	25	0.02	

TABLE 4.1 Comparison of different sample mounting methods to improve the throughput of SPIM systems.

4.3.2 SEMI-AUTOMATIC UNWRAPPING SOFTWARE

To allow the user to perform the image unwrapping and analysis during the screening process, I have created a **Qt** based Graphical user interface (GUI) that provides an integrated experience for the complex analysis task. The software is intended to be used with the integrated data pipeline described in **Chapter** 2. The GUI allows the user to check the quality of the image unwrapping software and requires very little user input.

4.3.2.0.1 Data import and ROI selection The first step of the image analysis workflow is data importing. After reading the full 3D-stack from the centralized file server, a maximum intensity projection of the image is displayed and allows users to select the region of interest for the analysis. It is important to select the part of the body that is suitable for the subsequent cylinder fitting step. The yolk extension of the embryo is the ideal body segment. The step also serves as an inspection step to see if the sample has reached the desired developmental stage for analysis (**Fig** 4.16).

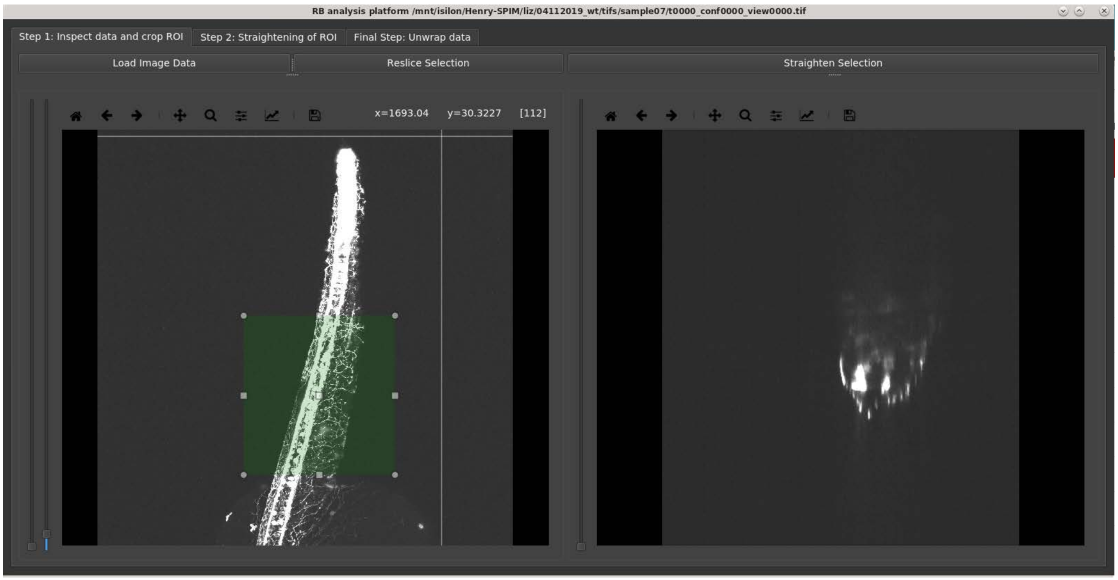


Figure 4.16 Step 1: Data loading and inspection screen. The left panel displays the MIP of the imported image stack (along with contrast control). After the user select the yolk extension ROI, the cropped image stack is displayed in the right panel.

4.3.2.0.2 Image straightening and cylinder fitting After the ROI selection process, the cropped image can then be straightened using the previously described body axis estimation method. This step is very robust and does not require any manual adjustment. After the image is straightened, the user can browse through the processed stack along the anterior-posterior axis. Optionally, users can crop the selected stack to accelerate the downstream processing further. Afterwards, an ellipse is fitted to each frame in the straightened image. This step requires the user to adjusting the smoothing parameters such to ensure the accuracy of the surface estimation (**Fig** 4.17). This is the only step in the GUI that has a tunable parameter. After the initial ellipse estimation, re-fitting the ellipse with a new smoothing parameter is fast.

4.3.2.0.3 Image unwrapping and data export After the ellipse fitting step, the image is ready to be unwrapped. The unwrapped image can be further cropped such that the RB cell body is not in the final image as they do not have a preferred orientation due to their circular shape. The final result can then be exported for directionality analysis in *Fiji* (**Fig** 4.18).

The overall runtime of the image analysis pipeline is around 1 to 1.5 minutes. It is worth noting that a significant portion of the runtime is spent reading the data into the software. We plan on developing the software further to allow a TCP/IP connection

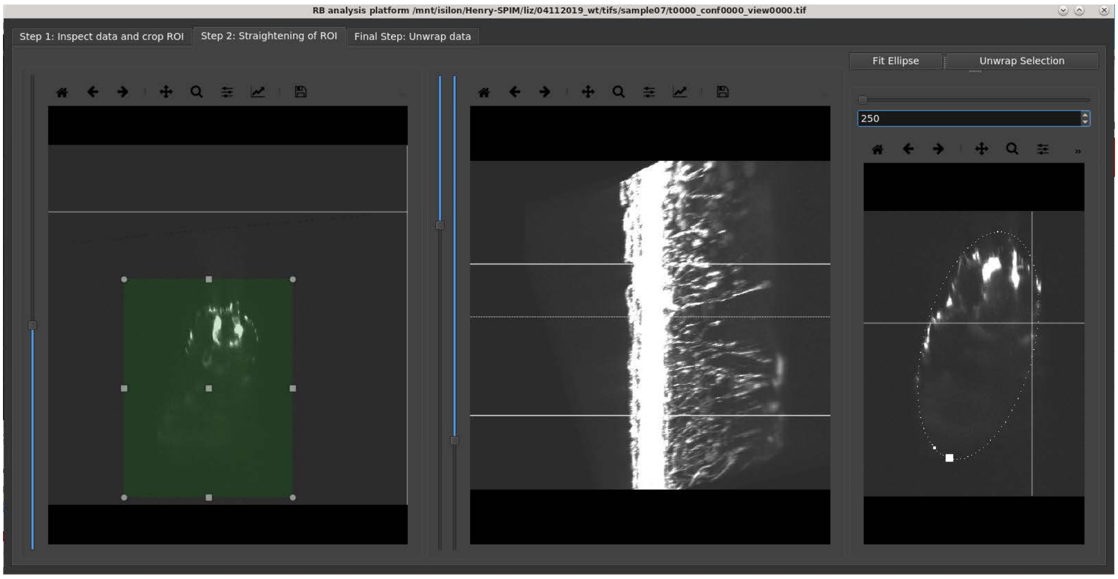


FIGURE 4.17 Step 2: Straightening inspection and ellipse fitting screen. The left panel show the straightened image stack. The middle panel displays the side view of the same stack. Users then need to select the range where the yolk extension is. The right channel is used to display the fitted ellipses and allow the user to fine-tune the fitting parameter.

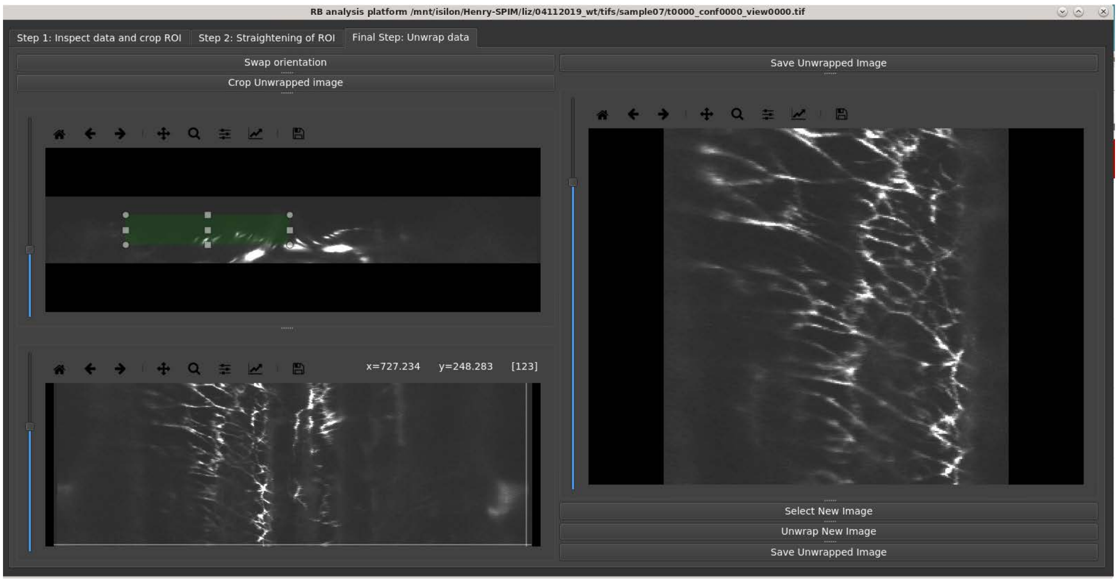


FIGURE 4.18 Step 3: Final inspection and output screen. In the final screen, user can inspect the unwrapped image stack and save the output to a *.tiff* file.

between the microscope and the image analysis software such that the data is analyzed in memory without having to go through the intermediate fileserver.

4.3.3 Discussion and outlook

I have described an integrated SPIM imaging and quantification workflow for zebrafish RB development imaging. I have verified that the combination of sample unwrapping and orientation analysis can be used to identify phenotypic differences in axon branching. The new integrated workflow has the potential to image up to 25 fish embryos with a single mounting procedure. More importantly, the images of the 25 samples can be analyzed with my custom analysis software right after acquisition.

In the next step, we will perform targeted genetic perturbation experiments to elucidate the underlying agents governing axon branching. The samples will be screened at a fixed developmental time point (24 hpf) when the RB axon projection growth becomes substantial. Currently, the orientation analysis step is performed outside of the unwrapping GUI. In the future, we plan on incorporating the *Fiji* processing steps into the pipeline so that the analysis result can be displayed along with the unwrapping result for a more integrated experience.

The screening system also represents another paradigm in the smart microscopy concept. With the quantification process built into the experimental procedure itself, the information-rich data product is output directly at the microscope. Therefore it is no longer necessary to save the overburdening amount of light sheet data. Only the quantification result need to be saved. Large SPIM raw images are only temporarily stored on the file server and the analysis machine. There is a significantly reduced long-term data storage burden.

It is worth noting that the quantification process is very tailored to the sample type. The method cannot easily be used in other image analysis applications. The concept of on-the-fly quantification often requires sacrifices in the complexity of algorithms used. As we are building a workflow for a phenotypic screening system with SPIM, addressing the data burden is crucial to the throughput of the system. Further verification needs to be performed to ensure the quantification process is robust and versatile.

One other concern regarding the image unwrapping method is the bias introduced by PSF anisotropy. Since the lateral resolution of the system is much higher than the axial resolution, image unwrapping would rotate the PSF to an arbitrary direction, potentially introducing smearing (visible in **Fig** 4.10). It might be necessary to deconvolve the image first before running the unwrapping software. Fortunately, Schmid and Huisken, 2015 have shown that with the help of GPU processing, deconvolution can be performed in real time.

The GUI currently requires some user input to perform ROI selections. These tasks can be automated with unsupervised pattern recognition approaches. Currently, the user input is saved into a log file, and I aim to train a convolutional neural network (CNN) to mimic user behaviors.

Once the unwrapping pipeline is fully automated, I plan on incorporating the temporal dynamics of the axon branching pattern into the phenotyping pipeline. Parameters such as axon growth rate and average branching length are key dynamics features that can be used to characterize phenotypic differences.

Although the image analysis suite is application specific, the high-throughput zebrafish screening system we are currently constructing can be used in a variety of applications. The imaging system is built upon the Flamingo sharable light sheet system project (https://involv3d.org/flamingo/), which will soon become available to the broader research community.

OTHER PROJECTS 5

> *"Knowledge is no burden."*
> *— Chinese Proverb*

SPIM has been widely accepted as the method of choice for long-term, low photo-damage fluorescence imaging. It has been used extensively in developmental biology and neurobiology prevalent model organisms such as zebrafish and fruit fly. Researchers trying to incorporate SPIM into other fields of study or other sample types, still face many challenges. In this chapter, I describe a few applications that I have created custom SPIM imaging protocols for, illustrating how the smart microscopy concept can help make SPIM a more approachable technique.

5.1 METABOLIC IMAGING OF PATIENT-DERIVED CANCER SPHEROIDS

This project demonstrates the application of SPIM in translational research. The high imaging speed of SPIM accelerates the screening process of cancer spheroids. Custom image analysis software facilitates quantification and hypothesis testing.

5.1.1 BACKGROUND

Spheroids are 3D complex cell culture that are more physiologically relevant than conventional cell culturing methods. Research has shown that spheroids resemble the metabolic activity and heterogeneity of the organs that the cell are derived from (Sharick et al., 2019). As a result, spheroids have been rapidly adapted to a wide variety of research fields. In cancer research, researchers have been able to use patient biopsy samples to grow spheroids of different cancer types (Drost and Clevers, 2018). Spheroids capture the same genetic abnormalities as their parent disease tissues and better recapitulate *in vivo* tumor characteristics than 2D cultures, which were widely used for research. A single patient biopsy sample can be regrown into hundreds of spheroids, which opens the door to parallel drug testing of standard-of-care therapies, drug combinations, and emerging treatments in a personalized fashion.

Imaging-based screening systems often are a tradeoff between content and throughput. High content methods, including laser-scanning microscopy, immunofluorescence, and RNAseq, are low-throughput or destructive. High-throughput approaches, such as epifluorescence microscopy are constraint by the limited metrics available to differentiate treatment responses. Researchers often use gross changes in spheroid diameter as a measure of drug effectiveness which is often inaccurate.

Another consideration for imaging-based screening systems is the metric for drug performances. Antibodies and dyes are often used to fluorescently label specific cellular structures such as the nucleus and mitochondria. However, the signal readout, in this case, does not intrinsically reflect the health of the spheroids when applying different drugs. Researchers have turned to utilize the autofluorescence signal from metabolic co-enzymes nicotinamide adenine dinucleotide (NAD) and Flavin adenine dinucleotide (FAD) to directly measure the metabolic activity of these spheroids and quantify their health (Alhallak et al., 2016; Heaster et al., 2016; Skala et al., 2007; Walsh et al., 2013). It has been shown that spheroids undergo metabolic shifts after treatment with standard-of-care and targeted therapies in many cancer types. The metabolic activity of the spheroids can be measured as the fluorescence signal ratio of the NADH and FAD, optical redox ratio:

$$R_{\mathrm{redox}} = \frac{I_{\mathrm{NADH}}}{I_{\mathrm{FAD}}} \tag{5.1}$$

To address the throughput problem with conventional screening system, we set out to create a patient-derived cancer organoid screening workflow. First, we performed mid-throughput metabolic imaging of cancer spheroids in a standard multi-view light sheet system as a proof-of-concept study.

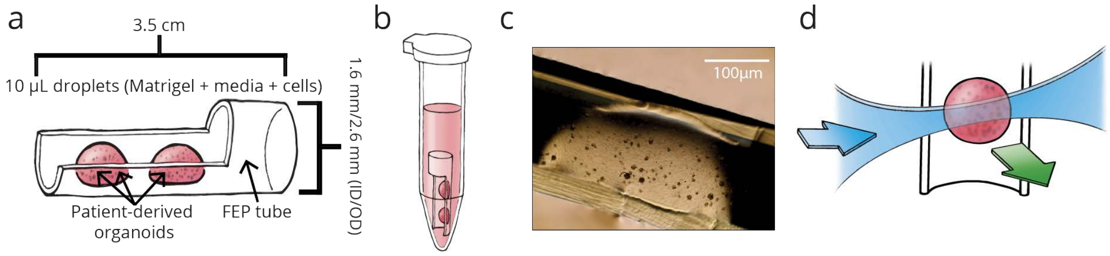


Figure 5.1 (a) Modified FEP tube with cut-out for easy spheroid culture in Matrigel droplets. (b) Spheroids were cultured in a 1.5 mL Eppendorf tube submerged in feeding media. (c) Brightfield image of spheroids growing in FEP tube. (d) Diagram of excitation (blue arrow) and emission (green arrow) illustrating selective plane illumination of Matrigel droplets. Illustrations by Jacki Whisenant.

5.1.2 Challenges

There are a few challenges imaging spheroids in a standard light sheet system. First of all, spheroids need to grow in a mixture of Matrigel and culture media. Since they are fragile, it is impossible to first grow them in a petri dish and then transfer to a standard FEP tube for imaging. Therefore, the spheroids have to be cultured inside a container that is compatible with the standard light sheet mounting assembly (See C.1.2).

Also, there is currently no image analysis software to perform the redox ratio quantification for a large light sheet image stack in a high throughput fashion. Cultured spheroids often exhibit a large variance in size. Larger spheroids tend to contain a prominent necrotic core with a bright FAD signal in the center. The bright signal can also be observed on the outside of the spheroid treated with drugs. This is largely due to the outer ring containing mostly dead or dying cells. With a higher imaging throughput, it becomes unfeasible to manually curate the image stacks to remove these functionally irrelevant layers of the cell. Therefore a custom automated analysis pipeline is needed.

To address these issues, we created a custom light sheet compatible organoid culturing protocol and an automatic shell based redox ratio analysis pipeline. This work is done in collaboration with *Dr. Peter Favreau* and *Dan Gil* from *Prof. Melissa Skala's* lab at UW-Madison.

5.1.3 Results

5.1.3.1 Custom sample preparation

In order to utilize the standard FEP tube holder used in most light sheet systems, we developed a custom sample preparation pipeline that uses modified FEP tubings and adapted existing spheroids culturing protocols: Spheroids were singularized with 0.25% Trypsin, mixed with Matrigel at a 1:1 ratio, and allowed to grow for 12 days in modified DMEM/F12 media supplemented with Wnt-conditioned media from murine L-cells cultured within modified FEP tubes. Modified FEP tubes were created from standard FEP tubes (1.6 mm inner diameter/2.6 mm outer diameter, ProLiquid) cut to lengths of 3.5 cm. These smaller tubes were cut in half lengthwise to allow Matrigel droplets to be placed inside the FEP tube and fully exposed to feeding media and treatments. Each FEP tube was placed in a 1.5 ml tube to submerge each Matrigel droplet with feeding media fully. A single Matrigel droplet could contain tens of spheroids (**Fig** 5.1 c), and multiple droplets could be placed in each tube (**Fig** 5.1 a).

This method allows us to mount the entire modified FEP tube into the standard SPIM system using the regular sample holder. Also, since the sample and matrigel culture is exposed to the liquid in the chamber, we can use the perfusion system in the microscope to exchange media and maintain the optimal temperature of 37 °C during experiments.

5.1.3.2 shell based image analysis

To remove the bias in redox ratio caused by the existence of apoptotic cells and the size variance between cultured spheroids, we developed a shell-based segmentation and analysis algorithm to quantify treatment response in spheroids as a function of depth relative to each spheroid's outer surface. This algorithm first performs 3D segmentation of spheroids within each SPIM volume using fuzzy C-means clustering and shape-based analysis (sphericity and size) to select for spheroids and exclude cell-debris. Each SPIM volume contains between 3 and 15 spheroids. The shell-based segmentation seeks to create hollow 3D surfaces (shells) with decreasing volume to quantify the pixel values in each shell. The 3D logical mask of a spheroid is first eroded using a spherical structure element and subtracted from the original 3D logical mask to yield the hollow 3D shell. This occurs iteratively until the algorithm reaches the center of the spheroid. These shells enables the quantification of NADH, FAD, and optical redox ratio as a function of depth. By rejecting the influence of the necrotic core

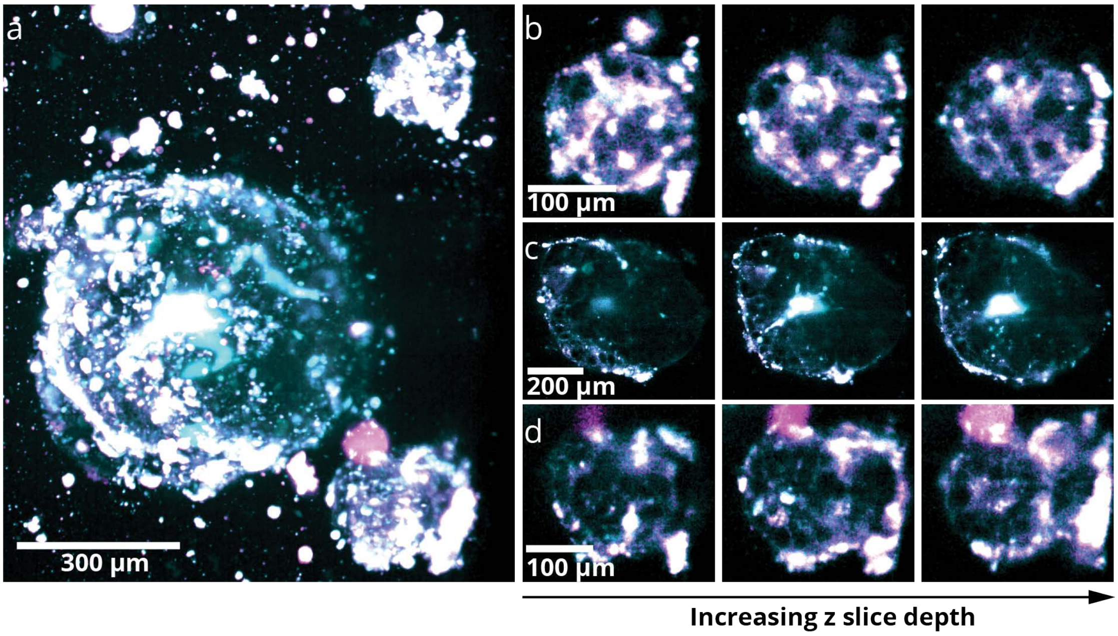


FIGURE 5.2 (a) Max intensity projection of three spheroids. (b-d) Selected optical slices of each spheroid. Magenta: NADH; Cyan:FAD.

and outermost dead cells, only the proliferating cells within the spheroids are used in measurements of treatment response (**Fig** 5.3).

5.1.3.3 COLORECTAL CANCER SPHEROIDS RESPONSE STUDY

To verify the validity of our workflow, we compared the redox ratio changes in metastatic colorectal cancer spheroids treated with different standard-of-care therapies. After 12 days of initial growth period, patient-derived spheroids were treated with media infused with different therapies for colorectal cancer, including either 10 μM 5-fluorouracil (5-FU), 40 μM oxaliplatin, or a combination of the two. Spheroids were then grown in treated media for an additional 48 hours before imaging experiment.

We first used the traditional 2D measurement of diameter as a metric for treatment response in spheroids as the ground truth. We compared both functional (redox ratio) and morphological (volume, surface area, and sphericity) features with the 2D diameter measurements to establish the strength of our approach. Our results indicate no significant change in volume or surface area compared to control after 48 hours of treatment, consistent with diameter measurements.

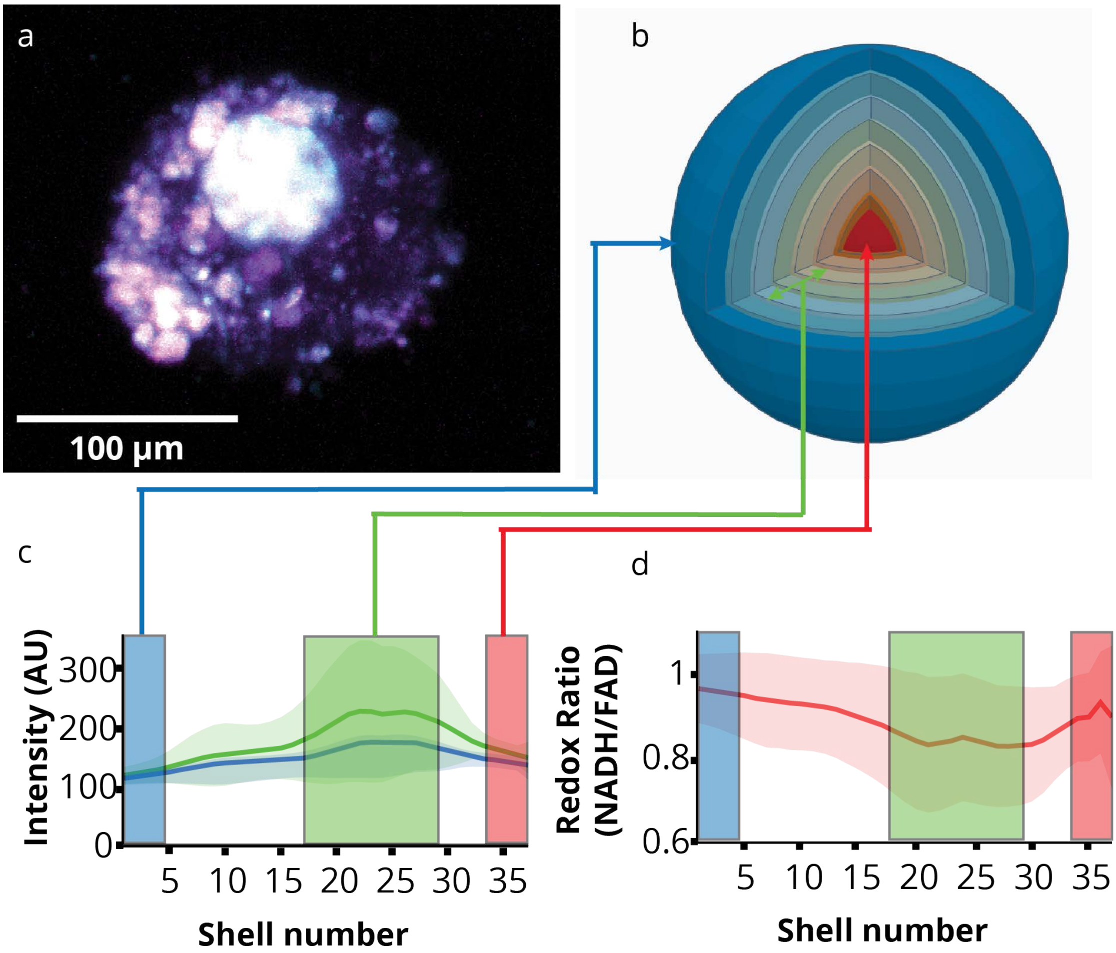


Figure 5.3 (a) Representative max intensity image of a colorectal spheroid with necrotic core in focus. Magenta: NADH; Cyan:FAD.(b) 3D spherical model of a typical spheroid. (c) NADH (blue) and FAD (green) intensity plots as a function of shell number. The blue arrow indicates the region corresponding to the outermost shell; the green arrow is a region comprising the proliferative zone; the red arrow represents the centermost or necrotic region of the spheroid. (d) The optical redox ratio as a function of shell number. Figure generated in collaboration with *Dan Gil*.

We next compared our metabolic results to previously measured multiphoton results of the same spheroid sample and treatments. Our results are consistent with the multiphoton measurements and indicate a highly significant response to Oxaliplatin and combination therapy of 5-FU and Oxaliplatin (**Fig** 5.4). The results are currently being prepared for publication.

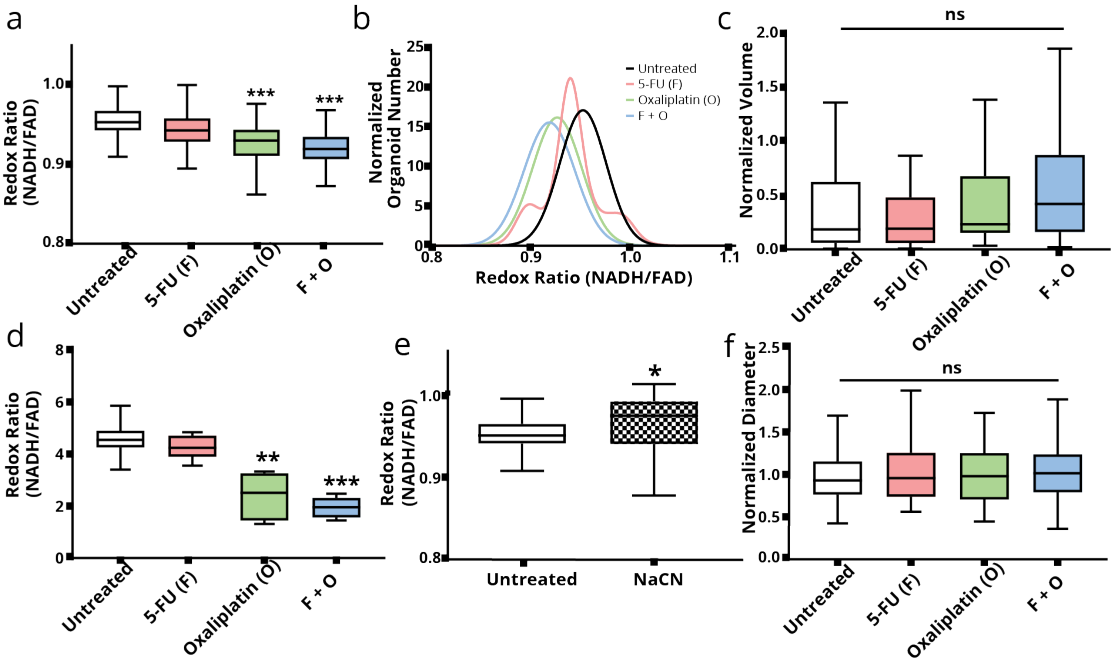


FIGURE 5.4 (a) Average optical redox ratio results from volumes acquired with SPIM and analyzed with a shell-based algorithm approach. (b) Subpopulation analysis of treatment response from spheroids separated by treatment. (c) Average volume results extracted from spheroid volumes with SPIM. (d) Average optical redox ratio results from images acquired with previously-studied multiphoton microscopy. (e) Optical redox ratio results of 10-mM NaCN control study acquired from SPIM. (f) 2D diameter measurements of CRC spheroids acquired from brightfield microscopy.

5.1.4 Discussions and outlooks

In this study, we established a new imaging pipeline to enable metabolic autofluorescence imaging of patient-derived spheroids in a standard SPIM system. With both the sample embedding and image processing issues addressed, we confirmed the validity of our pipeline in this proof-of-concept study. The sample embedding method is suitable for any SPIM systems using a vertical tube based mounting method and therefore can also be utilized with many commercial systems.

The next step is to increase the throughput of our system and integrate the image analysis pipeline into the imaging process similar to work done in **Chapter** 4. In the tube culturing method mentioned above, each FOV can contain up to 5 spheroids when using a 20× objective. Imaging the entire tube holder takes roughly 10 minutes due to manual FOV selection. The throughput is limited by the tube mounting method, as discussed in **Chapter** 4. We plan to use the same screening platform from **Chapter** 4 to image patient-derived spheroids. Although the culture is naturally grown in a petri dish environment, the depth at which the spheroids form remain arbitrary and therefore the tube embedding methods can still be used to constrain the location of growth. We also aim to build an automated image analysis pipeline similar to one described in **Chapter** 4. Image stacks captured by the system are directly analyzed, and only the shell based profiles are saved. This would also require implementing automated object recognition with CNN to eliminate the need for manual ROI selection.

5.2 Imaging cortical excitability in *Xenopus laevis* oocytes

In this project, I used SPIM to image cortical excitability in *Xenopus* oocytes. I adapted a custom sample mounting protocol to be able to precisely orient the sample. I also developed a custom image compression software to reduce the volume of the imaging dataset.

5.2.1 Background

Cytokinesis in animal cells is powered by a contractile ring of actin filaments (F-actin) and myosin-2. The formation of the contractile ring is directed by the small GTPase Rho, which is activated in a discrete zone at the equatorial cortex (Bement et al., 2005). Rho is regulated by guanine nucleotide exchange factors (GEFs) and GTPase activating proteins (GAPs) (Miller and Bement, 2008). When Rho is active, it interacts with a variety of downstream effector proteins that regulate cytokinesis. While the textbook model of cytokinesis depicts linear relationships between the regulators, it fails to capture the flexibility and robustness of cytokinesis that has been observed. A more recent model, termed "cortical excitability model", better explains these differences and is based upon nonlinear feedback loops between Rho, GEFs, GAPs, and effector proteins. During cortical excitability, the cell cortex generates traveling waves of Rho and F-actin activity that are eventually focused at the cytokinetic furrow. The wave propagates via positive feedback at the leading edge and negative feedback at the trailing edge, and their dynamics are representative of an excitable medium (Bement et al., 2015). However, questions remain as to what the key regulatory components in the dynamical system are and how cortical excitability form and evolve.

To answer these questions, our collaborators *Ani Michaud* and *Prof. Bill Bement* from UW-Madison image *Xenopus laevis* oocytes that are injected with fluorescent probes to visualize the dynamics of cortical excitability. Previously, the behavior has been

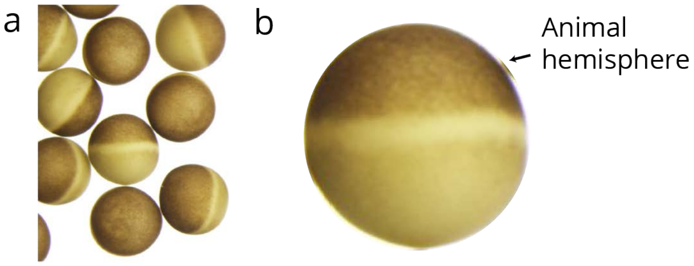


FIGURE 5.5 (a) Example image of a Xenopus oocyte. Image from Cristofori-Armstrong et al., 2015. (b) Xenopus oocyte is very opaque and we need to image the animal pole.

imaged with a confocal microscope with high spatial and temporal resolution (Bement et al., 2015). However, high laser dosage in a confocal microscope limits the duration of the imaging experiment. Besides, to achieve the necessary speed and resolution necessary to capture the behavior, confocal experiments sacrificed the field of view. Since the FOV has been smaller than the animal hemisphere of the oocyte, it has been difficult to capture the start of cortical excitability. For this reason, we set out to build a light sheet workflow to image the protein dynamics in *Xenopus* oocyte.

5.2.2 CHALLENGES

The Xenopus oocyte is optically very challenging to image. The opaque oocyte absorbs any excitation and fluorescence photons that pass through (**Fig** 5.5). The fluorescent signal comes exclusively from the surface of the animal hemisphere. Therefore it is critical to orient the sample such that the animal pole faces towards the detection objective such that neither the excitation photons nor the fluorescence photons need to travel through the inside of the oocyte. In a standard SPIM sample embedding protocol, the sample is taken into the FEP tube along with media. It is challenging to precisely orient the sample. Therefore we had to devise a new approach that allows manual re-orientation of the sample before imaging.

Another challenge that we encountered is the data storage required for each experiment. As we need to cover the entire oocyte surface, which spans more than 1 mm in diameter, a total of more than 250 images need to be taken in each channel. Typically, two protein factors are fluorescently labeled, which means two different channel images need to be taken. Also, due to the opaqueness of the sample, a second illumination arm is needed to cover the entire FOV. Overall, four image stacks totaling 1000 frames are taken at each time-point. Cortical excitability is highly dynamic. Hence, a

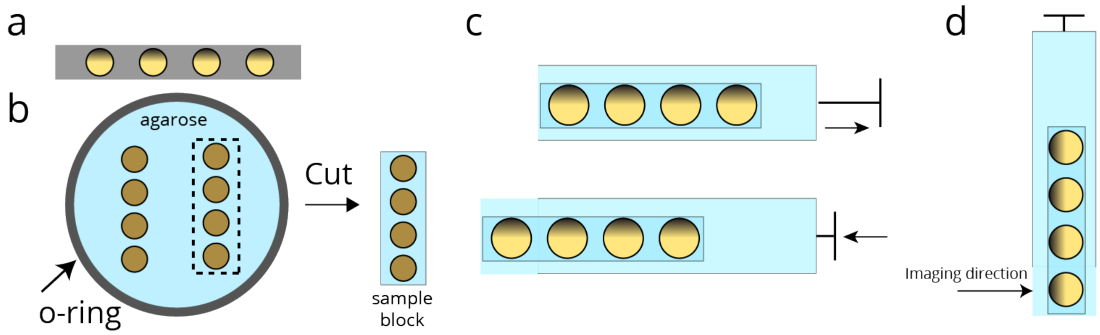


Figure 5.6 (a) Side view of oocytes embedded in a thin layer of agarose enclosed by an o-ring. (b) Top view of the o-ring pre-embedding step. Oocytes are oriented with a pipette tip such that the animal hemisphere is facing up. (c) The solidified agarose is then cut away from the O-ring and transferred into a syringe filled with agarose. The sample is then pushed out of the enclosure. (d) The oocyte now has the animal pole facing outward from the vertical column, ready for SPIM imaging.

full image stack needs to be taken every 20 seconds to be able to resolve the behavior temporally. To visualize the formation and evolution of cortical excitability over a typical 24 hours timelapse, approximately 30 TB of data need to be recorded. It is therefore critical to reduce the data size to relief the storage stress.

5.2.3 Results

5.2.3.1 Custom sample preparation

To address the sample orientation issue, we used an alternative sample mounting protocol utilizing a syringe. Instead of taking the sample inside the FEP tube with agarose, we first place the oocyte inside an o-ring that is filled with agarose. The oocytes are then oriented with a pipette tip such that the animal hemisphere faces upward. After the agarose solidifies, the agarose block containing oocytes is cut out of the o-ring and placed inside an agarose filled syringe. After the agarose solidifies inside the syringe, the sample are extruded out of the syringe for imaging (**Fig** 5.6).

5.2.3.2 Data size reduction

The amount of data generated in this experiment can easily overwhelm any standard lab storage solution. Hence, the data need to be compressed before permanent storage. In **Chapter** 2 we showed that the data could be effectively compressed down to about 50% the original size using integer compression methods, but the resultant data vol-

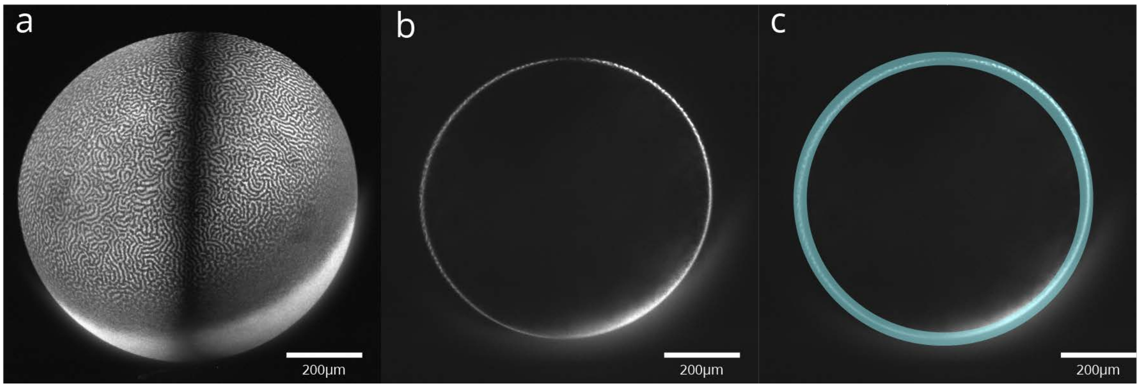


FIGURE 5.7 (a) Maximum intensity projection of a double-sided illumination stack. (b) A single slice of the image contains only a ring of foreground pixels(c) Fitting a circle to the image allows masking of the majority of the data.

ume is still large. Since the signal is strictly on the surface of the sample, the majority of the image volume contains no information. Similar to a previous study (Schmid et al., 2013), where a sphere is fitted to a zebrafish embryo, we can deploy a similar strategy to only save information that lies on a spherical shell. However, here it is not possible to estimate the spherical shell using the transmission image since the sample is opaque. Instead, we have to fit a circular shape to each image slice. The ring of fluorescence signal within the image represents the intersection between the light sheet and the spherical shell where the signal comes from (**Fig** 5.7 b). The thickness of the shell is effectively the illumination penetration depth of the light sheet which can be pre-determined (**Fig** 5.7,c). Instead of saving the entire image stacks containing mostly empty pixel, only the centroid, and radius of all the circles along with pixel values on the shell need to be saved. A 95% reduction in data size is achieved.

5.2.4 DISCUSSIONS AND OUTLOOKS

With both sample preparation and data storage issues addressed, we were able to capture detailed dynamics with two color labeling two different proteins (**Fig** 5.8). However, the data compression process cannot be performed on-the-fly due to the short interval between time-points. As a result, each experiment still requires around 30TB of temporary storage space before being compressed using the method described above. Without the help of dedicated hardware such as FPGAs and GPUs, we have to compress the data after the acquisition. It is also worth noting that even though the double-sided illumination improved the coverage of the sample, there is still a significant gap between the two views that potentially require one additional view to cover. It would be interesting to see if the smart rotation pipeline from **Chapter** 3 would be able to identify the optimal view combinations automatically.

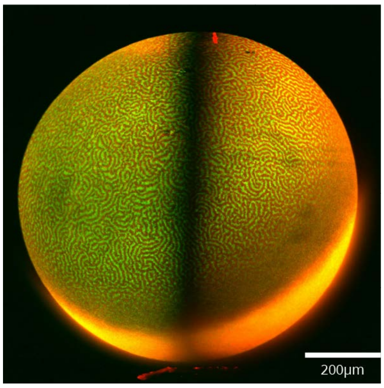


Figure 5.8 Sample still image from Oocyte experiment. Red: F-actin, Green: Active Rho.

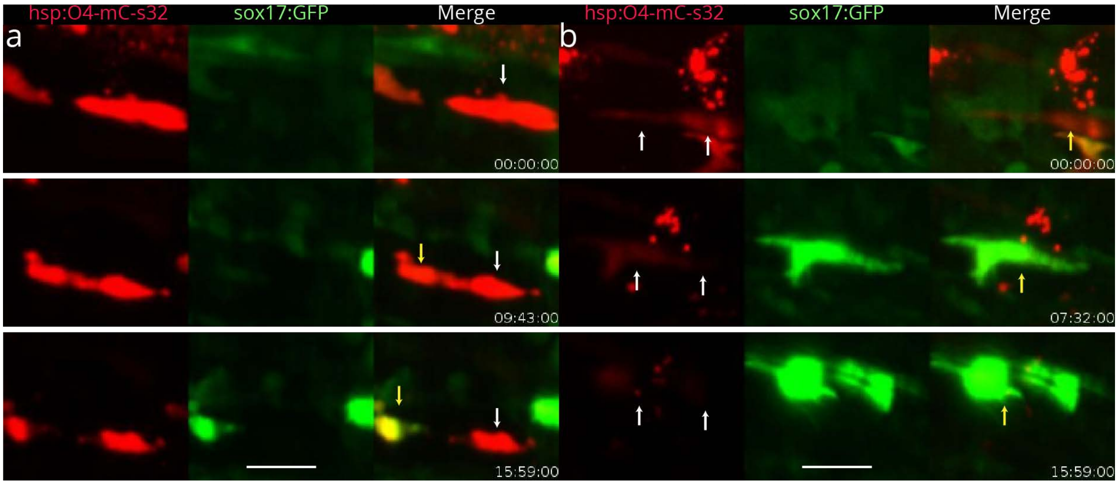


FIGURE 5.9 Lineage conversion *in vivo* imaged with SPIM.. Selected live image movie stills of individual myocytes with transgene expression (Red; *hsp:Oct4-P2A-mCherry-P2A-sox32*)in transgenic zebrafish (Green, *Tg(sox17:GFP)*) from 48-72hpf using SPIM. (a) Arrow point to an individual myocyte that splits into two presumptive cells in about 15 hours with the cell on the left (yellow arrow) exhibiting an increase insox17:GFP expression and adoptinga stellate shape. (b) Arrows point to a single cell rapidly changing color (red to green) and shape in less than 15 hours. Scale bar = 20 µm

5.3 MISCELLANEOUS PROJECTS

So far, in this thesis, I have discussed how different implementations of the smart microscopy concept can enable experimental workflows that were not possible otherwise. However, not all imaging experiments require complex quantification. The many advantages of SPIM over the more widely used confocal microscope can enable many experiments in its own right. During my thesis study, I had the pleasure to work with scientists with a wide variety of research interest. Here I will showcase a few collaborations where using SPIM has resulted in findings that were not possible before.

5.3.1 *In vivo* CELL LINEAGE CONVERSION IN ZEBRAFISH

In collaboration with *Prof. Duc Si Dong's* lab at Sanford Burnham Prebys Medical discovery Institute, we studied the reprogrammability of differentiated cells in their native microenvironment. It was believed that it is impossible for cells to be reprogrammed into cell types originating from a different germ layer. We found that ectopic co-expression of zebrafish *Sox32* with mouse *Oct4* in several non-endoderm lineages such as skeletal muscle cells can cell-autonomously trigger the early endoderm genetic program. Confocal microscopy is too photo-toxic for long term imaging to visualize

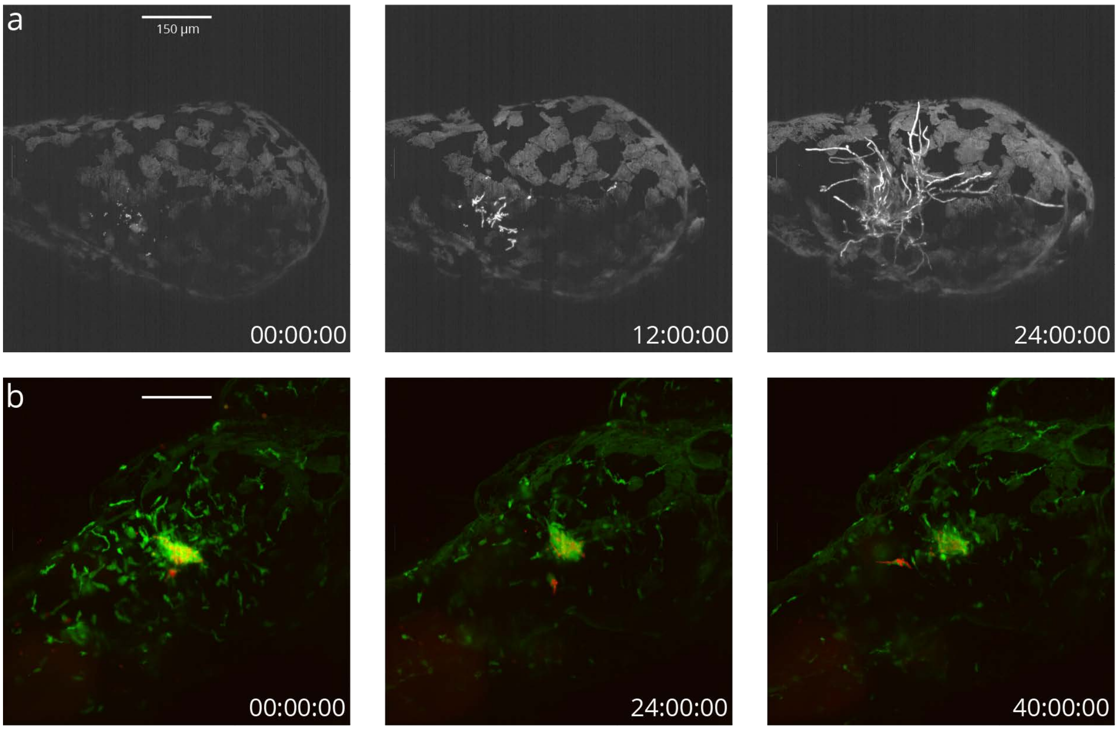


FIGURE 5.10 SPIM imaging of neutrophil-fungus interaction in zebrafish. (a) Zebrafish larvae that is neutrophil-defective (*mpx:rac2*) injected with RFP expressing *A.fumigatus*. The sample is treated with voriconazole but still shows continuous growth. (b) Zebrafish sample with labelled macrophages (Green, *Tg(mpeg1:EGFP)*) shows no substantial fungus growth (Red, *mpx:rac2*) in the presence of healthy neutrophils.

the cell fate. With the help of my SPIM system, we visualized the event where endoderm induced muscle cells to lose muscle morphology while gaining endoderm organogenesis markers over 24 hours (**Fig** 5.9).

5.3.2 IMAGING IMMUNE CELL - FUNGUS INTERACTION *in vivo*

Invasive fungal infection threatens the life of more than 2 million people globally each year, primarily in immunosuppressed patients (Brown et al., 2012). Most frontline drugs show significant effectiveness against life-threatening fungus such as *Aspergillus fumigatus in vitro*, yet patients with invasive aspergillosis who are treated only have a 50% survival rate (Baddley et al., 2010). This is largely due to the presence of immune cells and the complexity of whole tissues and organs affecting drug delivery and efficacy (Felton et al., 2014). To better understand the mechanism involved in the interaction between the antifungal drugs and host immune system, our collaborators

Prof. Emily Rosowski and *Prof. Anna Huttenlocher* used zebrafish as a model as it has been proven to be able to recapitulate the pathogenesis of invasive fungal infections (Rosowski et al., 2018). We performed SPIM imaging of zebrafish infected with *A. fumigatus* in the hindbrain with voriconazole treatment. We found that the presence of macrophages has a significant impact on the efficacy of the drug as macrophages slow down the progression of the disease to give the chemical more time to take effects *in vivo* (**Fig** 5.10). The long term imaging is only possible with SPIM as the sample health is severely compromised during the host infection, and a standard confocal time-lapse experiment was too damaging to allow long term observation.

5.3.3 NEUTROPHIL REPOLARIZATION UPON CONTACT WITH MACROPHAGES DURING WOUND RESPONSE

Reverse migration of neutrophils away from a wound site plays an important role in resolving neutrophilic inflammation; however, the mechanisms that regulate this process remains largely unknown. In a recent study, using larval zebrafish as an in vivo model of inflammation, it is shown that neutrophils engage in cell-to-cell contact with macrophages at the tail wound (Oliveira et al., 2016). This physical interaction subsequently induces neutrophil reverse migration. It is hypothesized that macrophages may promote neutrophil reverse migration by contact inhibition of locomotion (CIL). One key step in CIL is repolarization of colliding cells that involves a switch in rear-front polarity.

In collaboration with *Dr. Veronika Miskolci* and *Prof. Anna Huttenlocher*, we used high-speed multi-color SPIM imaging to test whether neutrophils repolarize upon macrophage contact. We used the calponin homology domain of utrophin tagged with GFP (*UtrCH-GFP*) that labels stable actin localized in the uropod of migrating

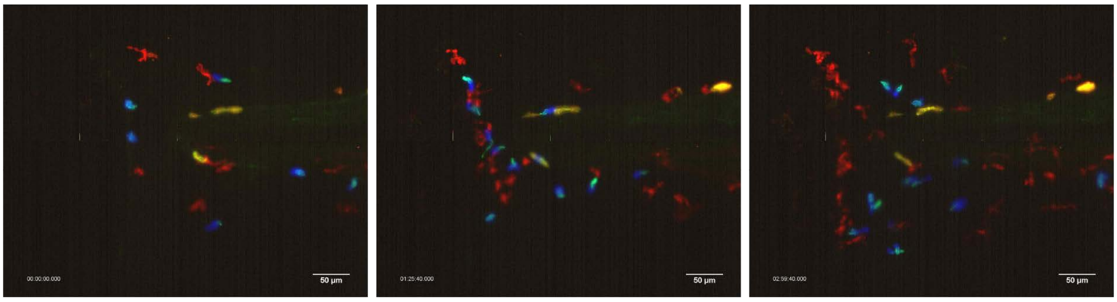


FIGURE 5.11 SPIM imaging of neutrophil repolarization during wound healing. Zebrafish embryo (*Tg(mpx:UtrCH-GFP)* × *Tg(lysC:BFP)* × *Tg(mpeg1:mCherry-CAAX)*) at 3 days post fertilization is wounded in the tail. Rapid recruitment of macrophages (in red) and neutrophils (in blue) can be observed over 3 hours, and the stable actin distribution (in green) can be visualized within the neutrophils.

neutrophils. We monitored neutrophil-macrophage interaction at the wound site for 3 hours at 20 seconds interval in 3 colors (**Fig** 5.11). Zebrafish embryos' tail is wounded with a razor blade at 3dpf and imaged immediately after for 2 hours to visualize the recruitment of macrophages and neutrophils. We are currently working on creating an image analysis software to isolate each neutrophil and analyze their stable actin distribution relative to the direction of travel.

Part III

DISCUSSION

6

Discussion and outlook

"Before you become too entranced with gorgeous gadgets and mesmerizing video displays, let me remind you that information is not knowledge, knowledge is not wisdom, and wisdom is not foresight. Each grows out of the other, and we need them all."
— *Arthur C. Clarke*

The big data issue in biology is often associated with large scale "omics" projects. For example, the 100, 000 Genome project hosts around 21 Petabytes of data (Caulfield et al., 2017, See online at : 100k genome project). However, these projects are often community-driven and have dedicated industrial data centers for housing and distribution. As a result, the data burden is shared amongst a community and does not significantly impact individual labs that want to utilize the resources. In comparison, bio-image data are usually very application specific and therefore, are handled by individual labs or institutions. During my thesis study, approximately 350 TB image data were generated. We can foresee that soon it would become common for labs to accumulate image data at a similar rate when imaging modalities such as SPIM become more common.

Besides, the microscopy development community is starting to realize the importance of sample throughput, i.e. "$N > NA$" or "sample size more important than resolution". Imaging more samples to generate statistics may be more important to biologists than pushing the spatial-temporal resolution limit of a specific instrument. Further breaking the resolution barriers may enable new discoveries, however, to reach meaningful biological conclusions, multiple repeats need to be performed. As a result, building mid- to high-throughput imaging systems such as the one described in **Chapter** 4 will become another focus of microscopy development in the near future. Increasing the sample throughput will further scale up the amount of data each experiment generates if no precautions are taken. Image data burden is becoming an imminent issue.

In this thesis, I focused on SPIM, which is one of the most data intensive modalities. I have explained that although SPIM raw image data are large, the final data product such as cell tracks is relatively small in volume. Large data size does not necessarily translate to high information content. It is therefore critical to find ways the raw data to data product conversion efficiency to relieve the data size pressure without compromising information quality. Three main approaches were detailed in this thesis.

1. Data size reduction. (See **Chapter** 2.)
2. Data quality improvement. (See **Chapter** 3.)
3. Direct data product conversion. (See **Chapter** 4.)

Each approach addresses the data burden in its specific ways and can be combined for additive effect.

The software tools I built in this thesis are very image modality and sample settings specific. One may question the applicability of these tools to a wider range of biological experiments. I would like to stress that it is difficult to create generalized tools for the smart microscopy concept. Each imaging experiment is designed to answer specific scientific questions and often involve the generation of new data types (new transgenic animals and fluorescent labels). A generalized smart microscope that is applicable to any sample and label combinations requires a significant amount of *a priori* knowledge, which is often not present. Therefore, instead of trying to create generalized tools that may sacrifice performance, I built tailored workflows for each experimental need that illustrates different ways of realizing the image analysis integrated smart microscopy concept.

In this chapter, I discuss how to utilize further the tools I have created. I also outline several challenges in the development of future smart microscopy tools. I end with a discussion about the prospect of artificial intelligence integrated smart microscopy.

6.1 Future works

6.1.1 Information content map for adaptive optics

Adaptive optics is a class of techniques that improves image quality by changing the optical configuration of the microscope during acquisition. To optimize the image quality, a quality metric has to be used. Previously, people have used simple metrics such as image brightness as the optimization target (Jesacher and Booth, 2011; Park et al., 2017). These optimization methods aim to model the aberration function and compensate accordingly using adjustable optical components such as deformable mirror devices and spatial light modulators. However, in large field-of-view applications, total image brightness comes from both the fluorophore density and aberrations. Ignoring fluorophore distribution may lead to the optical system searching for a more densely labeled region rather than compensating for optical artifacts. In **Chapter** 3, I formulated the information content map representation. The map captures both the optical accessibility and the underlying labeling distribution of the sample. Therefore, I am working on building an adaptive optics system that utilize the map.

Currently, I am working on building a SPIM system with adaptive aperture control. Although higher numerical aperture (NA) theoretically offers better theoretical resolution, the realized image resolution also depends on the sample properties. In a wide-field detection regime, higher NA can lead to increasing detection of scattered photons due to larger collection solid angle. I want to build a detection system for a standard SPIM that adaptively control the detection NA to reach an optimal balance between scattered photon rejection and system resolution based on the information content map representation.

6.1.2 Towards a fully automated phenotype screening system

In **Chapter** 4, I described a mid- to high-throughput phenotype screening system. I want to finish the construction of the V-SPIM and fully integrate the analysis tool. With the help of the custom sample mounting mold, the position of the samples are fixed on the Petri dish, eliminating the need for manual sample positioning. The software analysis step is currently only semi-automatic. Users need to select the optimal ROIs for processing manually. The ROI selection process can be considered as a semantic segmentation step where the foreground area is separated from the background. Although conventional methods such as watershed segmentation are fast enough for on-the-fly operations, they usually require parameter fine-tuning to achieve optimal performance. Deep learning based semantic segmentation methods such as U-net, therefore, are parameters-free (Ronneberger et al., 2015). Currently, the quantification software record the user manual selected ROIs. The user behavior log can be used to generate a rough segmentation ground truth to train a segmentation network.

In addition, I am working on rewriting the shell based quantification software for the spheroid screening projects so that it is fast enough for on-the-fly analysis. The goal is to create an automated SPIM based spheroid screening system for personalized cancer therapy and drug screening applications.

6.1.3 Integrated smart microscopy platform

Although the specific implementation of the smart microscopy concept varies from applications to applications, the central idea of image analysis integrated microscopy always requires a dedicated data pipeline. The bidirectional data communication between the microscope, analysis engine, and data storage required are essential. I have built the "Huisken lab command listener" framework that runs on the analysis engine. The analysis engine becomes the master controller of the data flow. The software can only facilitate local TCP/IP connections as well as the execution of *Fiji* or *Python* scripts via a command line interface. In the future, I would like to extend the programming language support for the software to improve accessibility. I would also like to make the software compatible with **KNIME** (Berthold et al., 2007, The Konstanz Information Miner) platform to allow easier image analysis pipeline creation.

6.2 Challenges

There are still many challenges regarding the future development of the smart microscopy concept. During my thesis study, I always have to compromise between image analysis accuracy and speed. For example, in the smart rotation workflow described in **Chapter** 3, the information content map has to be summarized by angle, and only the 2D projections of the maps are compared. This is due to a lack of accurate and fast 3D registration methods that can efficiently register multiple 3D image stacks in a reasonable time. Other researchers have built GPU accelerated registration software to perform real-time registration (e.g. the Clear control framework https://github.com/ClearControl). The processing time for more than four views is still long. The fastest processing speed can be achieved by running the image analysis on FPGA. However, FPGA with sufficiently large memory to host microscopy images are often costly. Also, writing complex algorithms for FPGA is very time-consuming.

In addition, there is a disconnect between the bio-image analysis community and the general computer vision community. The computer vision community creates software using either *C++* for speed or *MATLAB* for user-friendliness. On the other hand, the bio-image analysis community mostly writes software in *Java* so that it can be easily integrated into the *Fiji* ecosystem. As a result, I had to rewrite much existing software in different languages to reduce the number of interpretation layers. It is also worth noting that most off-the-shelf microscopy hardware such as cameras and lasers have control software library in *C++* or *MATLAB*. Therefore, writing the image analysis software in these two programming languages can further enhance the integration between hardware and software.

In my opinion, one way to address both issues is for the camera manufacturers to integrate certain image processing algorithms into the camera. Many time consuming pre-processing steps such as Gaussian filtering can be performed efficiently on the FPGA of the camera, relieving pressure for downstream processing. Similar concepts have been implemented in industrial "intelligent cameras" where processing steps such as optical flow and template matching are performed in-camera.

6.3 AI FOR MICROSCOPY

In recent years, there has been an explosion of artificial intelligence (AI) applications in both academia and general public. For the first time, machine learning can perform complex tasks such as object recognition close to or exceeding human precision (Krizhevsky et al., 2012). Thanks to the advancement in commercial GPU processing power and availability of well curated data, neural network based AI technology is already changing many aspects of society including research. Although the training time of a neural network can be long, applying neural network to perform tasks such as object recognition is very fast. As a result, neural networks are rapidly replacing humans in many tasks where existing ground-truth samples are abundant.

There have been many studies that apply neural networks to microscopy images for different purposes. Most notably, Weigert et al., 2018 detailed the use of a U-net structure to improve the image quality of microscopy images. Ounkomol et al., 2018 used a similar method to generate fluorescence images from unlabeled transmitted-light images. It is very tempting to reach the conclusion that AI is going to resolve many challenges in smart microscopy development. However, like any other machine learning approache, the performance of neural network based AI algorithms depends on the quantity and quality of training data. Depending on the specific analysis requirement, there may not be a reliable way to generate accurate ground-truth datasets for training. Since existing image analysis tools are often limited in accuracy or speed, many researchers have created semi-automatic tools to manually annotate and curate SPIM datasets (Wolff et al., 2018). With a lack of training datasets for complex image analysis tasks, application of AI to microscopy images is limited to low-level processing tasks such as denoising (Krull et al., 2018).

Even with a significant amount of ground-truth, due to the challenging size of SPIM images, the amount of computational power required to train a neural network is high. Therefore, utilizing AI for SPIM image analysis is not currently practical. That is not to say that AI cannot contribute to the smart microscopy concept at all. One other area where AI is making significant contribution is autonomous machine operation. For example, although it is possible to use adaptive optics to correct sample aberrations, the process of optimizing the system configuration is very time consuming. Since it is possible to generate images with or without adaptive optics correction, AI has the potential to greatly accelerate the process of optimizing the optical configuration. In addition, much like self-driving cars, I believe the future generations of smart microscope can be fully autonomous. With the help of AI, the microscope itself can identify interesting image pattern based on previous user experience and record optimal image data at an adaptive sampling rate to further reduce the information redundancy, as envisioned in Scherf and Huisken, 2015.

Part IV

APPENDIX

SUPPLEMENTARY TO Chapter 2

A

The details of the three datasets used for data compression benchmarking is listed below:

- Dataset *Emily*
  Image data collected for the fungus immune system interaction project (Details in **Chapter** 5.3.2).
  Images were taken at 100 fps using the *Andor Zyla 4.2 PLUS* with 10 ms exposure time.
  100 image stacks were randomly selected from 5 different experiments. 50 image stacks from each channel.
  Each stack contains 400 to 500 full frame (2048 × 2048) images.

- Dataset *ani*
  Image data collected for the Xenopus oocytes cortical excitability project (Details in **Chapter** 5.2).
  Images were taken at 100 fps using the *Hamamatsu Orca Flash V2* with 10 ms exposure time.
  100 image stacks were randomly selected from 5 different experiments. 50 image stacks from each channel.
  Each stack contains 300 to 400 full frame (2048 × 2048) images.

- Dataset *liz*
  Image data collected for the zebrafish RB cells axon imaging project (Details in **Chapter** 4).
  Images were taken at 100fps using the *Andor Zyla 4.2 PLUS* camera with 10 ms exposure time.
  100 image stacks were randomly selected from 10 different experiments.
  Each stack contains 200 to 300 full frame (2048 × 2048) images.

SAMPLE HANDLING

B

B.1 ZEBRAFISH SAMPLES

Zebrafish samples, both adults and embryos, are handled according to established protocols and community guidelines (Nüsslein-Volhard and Dahm, 2001). All fish sample handling is according to EU Directive 2019/63/EU and German Animal Welfare Act for experiments performed in Dresden, Germany and UW-Madison Institutional Animal Care and Use Committee (IACUC) for experiments performed in Madison, USA.

B.1.1 TRANSGENIC ZEBRAFISH LINES USED

The transgenic zebrafish lines used in different parts of the thesis are listed in B.1. The

Transgenic line	Labelled structure	Reference	Location in text
Tg(h2afva:h2afva-mCherry)	Nuclei	Knopf et al., 2011	**Chapter** 3
Tg(kdrl:GFP)	Vasculature	Jin et al., 2005	**Chapter** 3
Tg(ngn1:GFP-CAAX)	Neurons	Blader et al., 2003	**Chapter** 4
Tg(sox17:EGFP)	Endoderms	Mizoguchi et al., 2008	**Chapter** 5.3.1
Tg(mpeg1:EGFP)	Macrophages	Ellett et al., 2011	**Chapter** 5.3.2
Tg(mpx:UtrCH-GFP)	Stable actin	Yoo et al., 2010	**Chapter** 5.3.3
Tg(lysC:BFP)	Neutrophil	Unpublished	**Chapter** 5.3.3
Tg(mpeg1:mCherry-CAAX)	Macrophages	Unpublished	**Chapter** 5.3.3

TABLE B.1 Zebrafish transgenic line used in this thesis unpublished fish lines used in **Chapter** 5.3.3 were generated by *Dr. Veronika Miskolci* from *Prof. Anna Huttenlocher* lab at UW-Madison.

B.1.2 ZEBRAFISH HUSBANDRY

Adult zebrafish are kept in the dedicated fish facility at the Wisconsin Institutes for Discovery. For mating and egg collection, a pair of female and male zebrafish are put together in a separate mating box in late afternoon. The mating box has a separate bottom compartment separated from the top with a mesh such that the eggs can fall through. Eggs are collected in the morning and washed with egg water. Eggs are then transferred into a petri dish with E3 medium and methylene blue and kept at 28.5 degrees Celsius until imaging. Zebrafish husbandry is performed either by me or Alyssa Graves from the Huisken lab at the Morgridge Institute for Research.

B.1.3 SAMPLE PREPARATION FOR IMAGING

Different variations of the sample mounting protocol detailed in Kaufmann et al., 2012 were used to prepare zebrafish embryos for SPIM imaging. The samples were mounted in an FEP tube of varying sizes and with low melting agarose of varying concentrations.

B.1.3.0.1 ***Chapter*** *3* Samples are embedded in FEP tubes (0.8mm inner diameter , 1.2mm outer diameter, Bola) with 0.8% low melting point agarose (*Sigma-Aldrich A9414*) made with E3 containing 200 mg per l Tricaine (*Western Chemical Inc.*). To demonstrate fusion result, fluorescent beads (Fluoresbrite Plain YG 0.5micron microspheres, Polyscience) are also mixed in the embedding agarose gel using a previously published protocol (Preibisch et al., 2010).

B.1.3.0.2 ***Chapter*** *4* Samples are taken into the FEP tubes (0.8mm inner diameter, 1.2mm outer diameter) along with standard E3 media with tricaine. The bottom of the FEP tube is then sealed with 1.2% agarose. For drug treatment studies, the 5μM *CK666* (*Sigma-Aldrich*) is added to E3 along with 2% DMSO. The media in the sample chamber's perfusion system is always matched with the media inside the FEP tube.

B.1.3.0.3 ***Chapter*** *5.3.1* 1.0nl plasmid is injected at the 1-cell stage to deliver 30 to 35ng of plasmid DNA. Embryos then undergo 5 cycles of heat shock from 24hpf to 48hpf (3 hours heat shock at 38.5 degrees and then 2.5 hours at 28.5 degrees). Samples are then embedded with 0.6% agarose in a FEP tube (0.8mm inner diameter, 1.2mm outer diameter) .

B.1.3.0.4 ***Chapter*** *5.3.2* To prevent pigment formation, 0.2mM N-pheylthiourea (PTU, *Sigma-Aldrich*) was added to E3 water at 12 or 24 hpf. Samples are embedded with 0.8% agarose in an FEP tube (1.6mm inner diameter, 2.4mm outer diameter). $1\mu g/ml$ voriconazole (*Sigma-Aldrich*) is also added to the agarose. The increased concentration is needed to constrain the embryo as increased involuntary twitches as a result of the fungal infection can be observed.

B.1.3.0.5 ***Chapter*** *5.3.3* Samples were embedded in an FEP tube (0.8mm inner diameter, 1.2mm outer diameter) with 0.6% agarose made with E3 and tricaine. Since the sample is imaged at relatively high frequency, higher concentration of agarose is used to increase the sample stability. No significant sample defects can be observed since the imaging duration is relatively short (less than 3 hours).

B.2 OTHER SAMPLES

B.2.1 PATIENT DERIVED CANCER SPHEROID

The sample preparation protocol is detailed in **Chapter** 5.1.3.1. Patient sample collection is performed in collaboration with Prof. Dustin Deming at UW-Madison. Spheroid re-culturing is performed by Dr. Peter Favreau at Morgridge Institute for Reserach. I performed the tube preparation, mounting and imaging with SPIM.

B.2.2 *Xenopus Laevis* OOCYTE

Xenopus wild type sample is injected with fluorescence probe by *Ani Michaud* from *Prof. Bill Bement* lab at UW-Madison. I then embed the sample according to the method detailed in **Chapter** 5.2.3.1.

C

HARDWARE

C.1 MULTI-VIEW LIGHT SHEET FLUORESCENCE MICROSCOPE

C.1.1 OPTICAL LAYOUT

The uncollimated laser from the laser source (LS) is collimated with an aspheric lens (COL) to obtain the required beam diameter. Collimated beam is then folded by mirrors (M0/FM) and passed through a periscope (PS/PS′) to get to the desired beam height. The bottom periscope mirror is a resonant mirror (RM/RM′) to perform destriping. The beam is then passed through a cylindrical lens (CL/CL′) for shaping. The reshaped beam is then passed through a telescope (L1/L2) and focused into the imaging chamber with an illumination objective (ILL/ILL′). The fluorescence from the sample is collected by a detection objective (DET) and focused onto the camera (CAM) sensor via a tube lens (TL). See **Fig** C.2 for detailed illustration.

C.1.2 SAMPLE HOLDER ASSEMBLY

The sample is usually held vertically within a FEP tube. The tube is then in turn held by a custom sample holder. The sample holder is placed onto a stage assembly so that the sample chamber along with the sample can be translated in 3D for sample finding and image generation. The platform where the sample rests on is also a high-precision rotational stage. The sample can be rotated around the vertical axis.

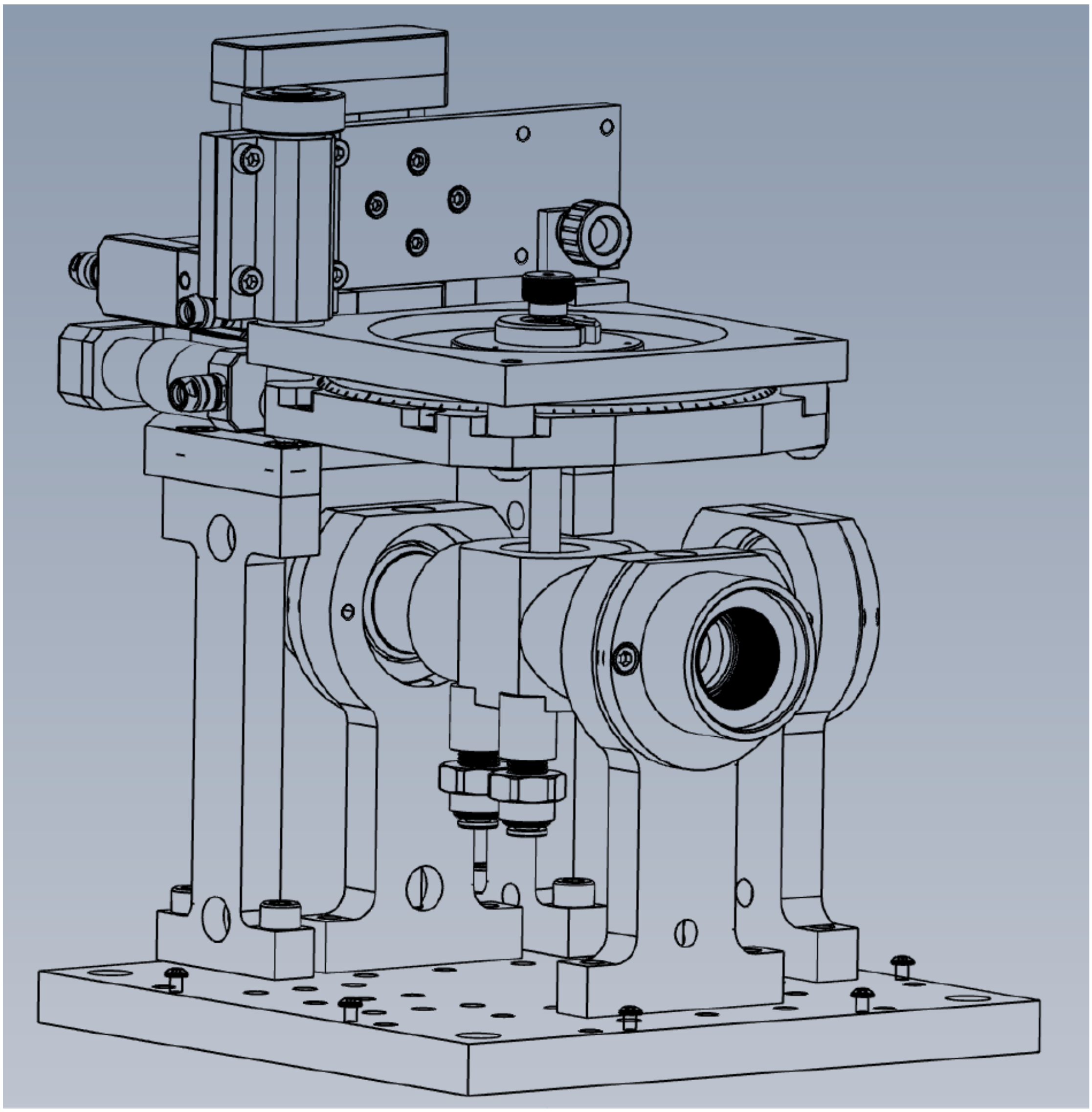


Figure C.1 The sample holder translation platform that is placed above the imaging chamber. Designed in collaboration with *Robert Swader* and *Todd Bakken* from Morgridge Institute for Research.

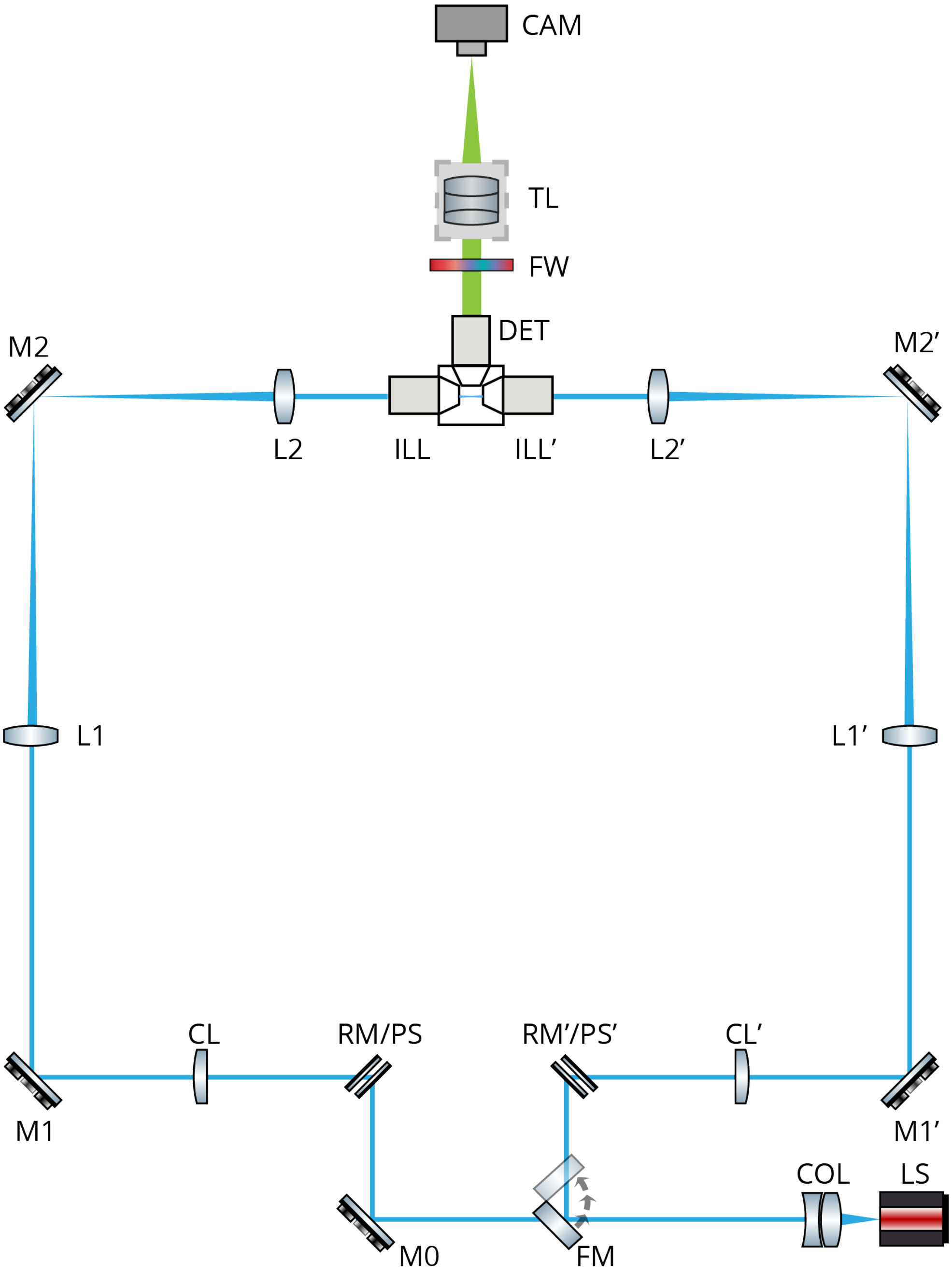


Figure C.2 The optical layout of the multi-view SPIM used in this thesis

C.1.3 Component list

Part name	Quantity	Vendor	Product ID	Code
Camera	1	Andor	Zyla 4.2 PLUS	CAM
Camera	1	Hamamatsu	Orca Flash V2	CAM
Laser engine	1	Toptica	MLE	LS
Illumination optics				
Illumination objective	2	Olympus	UMPLFLN10XW	ILL
100mm lens	2	Thorlabs	AC254-100-A	L2
200mm lens	2	Thorlabs	AC254-200-A	L1
50mm cylindrical lens	2	Thorlabs	ACY254-050-A	CL
Flip mirror	1	Newport	8892-K-M	FM
Mirror mount	8	Newport	U100-G2K	M0-M2
Goniometers	2	Thorlabs	GN1/M	accessory
Lens mount	4	Newport	LH-1A	accessory
Cylindrical lens mount	2	Newport	CYH-1	accessory
Aspheric lens	1	Newport	KPA19AR.14	COL
Detection optics				
Detection objective	1	Olympus	UMPLFLN10XW	DET
Detection objective	1	Olympus	UMPLFLN20XW	DET
Detection objective	1	Olympus	LUMPLFLN40XW	DET
Filter wheel	1	Ludl	96A351	FW
GFP filter	1	Chroma	ET525/50	accessory
RFP filter	1	Chroma	ET570lp	accessory
BFP filter	1	Chroma	ET450/50m	accessory
Far red filter	1	Chroma	ET655lp	accessory
Tube lens	1	Olympus	U-TLU	TL
Miscellaneous optical mechanics				
Optical posts	2	Newport	M-PS-3-PK	Misc.
translational stages	3	PI	M-111.1DG	Misc.
Rotational stage (optional)	1	PI	U-651	Misc.
stage controller	1	PI	C884	Misc.
stage controller	1	PI	C863	Misc.
Optical Rail	1	Owis	20′	Misc.
Rail platform	2	Owis	dove tail platform	Misc.

TABLE C.1 The list of components used to construct the custom multi-view SPIM system

Two different cameras were used during the thesis work: *Hamamatsu Orca Flash V2* and *Andor Zyla 4.2 PLUS*. There is no practical differences between the two as they operate at the same speed and quantum efficiency.

There were three different detection objectives used during the thesis work. The one with the most appropriate magnification level (10×, 20× or 40×) is chosen for each project. It is worth noting that the three objectives have the same external dimension. The same mounting mechanism as well as image chamber can be used for all three objectives which is a very convenient feature of my microscope.

C.2 INTEGRATED DATA PIPELINE

The image data collected at Morgridge Institute for Research is stored on the Huisken lab dedicated file server. The image analysis tasks (both on-the-fly and post-processing) are performed on a dedicated Linux compute server. The file server, Linux compute server and the microscope control computer are integrated via high-speed network according to the outline in **Chapter** 2.

C.2.1 MICROSCOPE CONTROL COMPUTER

The detailed specification for the microscope control computer:

- **Operating system**: Windows Server 2016
- **CPU**: Intel Xeon CPU E5-2630 v3 @ 2.40GHz ×2
- **GPU**: NVIDIA Quadro M5000
- **RAM**: 64GB
- **System drives**: 256 GB SSD ×2 in RAID 0
- **Storage drives**: 1 TB HDD ×2 in RAID 0
- **Network card** Mellanox ConnectX-4 Lx @ 25Gbps

C.2.2 CENTRALIZED STORAGE SERVER

The detailed specification for the Huisken lab storage server:

C.2.2.0.1 Performance tier The performance tier storage is used to store the newly captured data from the microscope. It is optimized for speed rather than volume. Data stored on the performance tier is automatically transported to the archive tier if it has not been accessed in 10 days.

- **Chasis**:Isilon F800
- **RAM**:1024 GB
- **Storage drives**: 1.6 TB SSD ×60
- **Network connection**: 40Gbps ×4

C.2.2.0.2 Archive tier Data that is idle for more than 10 days on the performance tier is automatically transferred to the archive tier. The archive tier servers are optimized for capacity and cost efficiency.

- **Chasis**:Isilon A200 ×2
- **RAM**:64 GB
- **Storage drives**: 8 TB HDD ×60
- **Network connection**: 10Gbps ×8

C.2.3 LINUX COMPUTE SERVER

The detailed specification for the Huisken lab Linux compute server:

- **Operating system**: CentOS 7
- **CPU**: Intel Xeon CPU E5-2698 v4 @ 2.20GHz ×2
- **GPU**: NVIDIA Quadro P5000
- **RAM**: 256GB
- **System drives**: 40 TB SSD drives
- **Network card** Mellanox ConnectX-4 Lx @ 40Gbps

The Linux server is kept close to the fileserver. Remote sessions are set up for each lab members and the computing resource is equally distributed amongst users.

BIBLIOGRAPHY

Alhallak, K., L. G. Rebello, T. J. Muldoon, K. P. Quinn, and N. Rajaram (2016). "Optical redox ratio identifies metastatic potential-dependent changes in breast cancer cell metabolism." eng. In: *Biomedical optics express* 7.11, pp. 4364–4374. DOI: `10.1364/BOE.7.004364` (cit. on p. 74).

Amat, F., B. Höckendorf, Y. Wan, W. C. Lemon, K. McDole, and (2015). "Efficient processing and analysis of large-scale light-sheet microscopy data". In: *Nature Protocols* 10.11, pp. 1679–1696. DOI: `10.1038/nprot.2015.111` (cit. on p. 20).

Andersen, E. F., N. S. Asuri, and M. C. Halloran (2011). "In vivo imaging of cell behaviors and F-actin reveals LIM-HD transcription factor regulation of peripheral versus central sensory axon development." eng. In: *Neural development* 6, p. 27. DOI: `10.1186/1749-8104-6-27` (cit. on p. 53).

Baddley, J. W. et al. (2010). "Factors associated with mortality in transplant patients with invasive aspergillosis." eng. In: *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 50.12, pp. 1559–1567. DOI: `10.1086/652768` (cit. on p. 86).

Balázs, B., J. Deschamps, M. Albert, J. Ries, and L. Hufnagel (2017). "A real-time compression library for microscopy images". In: *bioRxiv*. DOI: `10.1101/164624`. eprint: `https://www.biorxiv.org/content/early/2017/07/21/164624.full.pdf` (cit. on p. 20).

Bement, W. M., H. A. Benink, and G. von Dassow (2005). "A microtubule-dependent zone of active RhoA during cleavage plane specification." eng. In: *The Journal of cell biology* 170.1, pp. 91–101. DOI: `10.1083/jcb.200501131` (cit. on p. 80).

Bement, W. M. et al. (2015). "Activator–inhibitor coupling between Rho signalling and actin assembly makes the cell cortex an excitable medium". In: *Nature Cell Biology* 17, p. 1471 (cit. on pp. 80, 81).

Berthold, M. R. et al. (2007). "KNIME: The Konstanz Information Miner". In: *Studies in Classification, Data Analysis, and Knowledge Organization (GfKL 2007)*. Springer (cit. on p. 94).

Betzig, E. et al. (2006). "Imaging Intracellular Fluorescent Proteins at Nanometer Resolution". In: *Science* 313.5793, pp. 1642–1645. DOI: `10.1126/science.1127344`. eprint: `https://science.sciencemag.org/content/313/5793/1642.full.pdf` (cit. on p. 4).

Blader, P., C. Plessy, and U. Strahle (2003). "Multiple regulatory elements with spatially and temporally distinct activities control neurogenin1 expression in primary neurons of the zebrafish embryo." eng. In: *Mechanisms of development* 120.2, pp. 211–218 (cit. on p. 101).

Bria, A., G. Iannello, L. Onofri, and H. Peng (2016). "TeraFly: real-time three-dimensional visualization and annotation of terabytes of multidimensional volumetric images". In: *Nature Methods* 13, p. 192 (cit. on p. 11).

Brown, G. D., D. W. Denning, N. A. R. Gow, S. M. Levitz, M. G. Netea, and T. C. White (2012). "Hidden killers: human fungal infections." eng. In: *Science translational medicine* 4.165, 165rv13. DOI: `10.1126/scitranslmed.3004404` (cit. on p. 86).

Carpenter, A. E. et al. (2006). "CellProfiler: image analysis software for identifying and quantifying cell phenotypes". In: *Genome Biology* 7.10, R100. DOI: `10.1186/gb-2006-7-10-r100` (cit. on p. 5).

Caulfield, M. et al. (2017). *The 100,000 Genomes Project Protocol*. DOI: `10.6084/m9.figshare.4530893.v2` (cit. on p. 91).

Cheeseman, B. L., U. Günther, K. Gonciarz, M. Susik, and I. F. Sbalzarini (2018). "Adaptive particle representation of fluorescence microscopy images". In: *Nature Communications* 9.1, p. 5160. DOI: `10.1038/s41467-018-07390-9` (cit. on p. 21).

Chen, B.-C. et al. (2014). "Lattice light-sheet microscopy: imaging molecules to embryos at high spatiotemporal resolution." In: *Science* 346.6208, p. 1257998. DOI: `10.1126/science.1257998` (cit. on p. 11).

Chen, F., P. W. Tillberg, and E. S. Boyden (2015). "Expansion microscopy". In: *Science* 347.6221, pp. 543–548. DOI: `10.1126/science.1260088`. eprint: `https://science.sciencemag.org/content/347/6221/543.full.pdf` (cit. on p. 4).

Chen, K. H., A. N. Boettiger, J. R. Moffitt, S. Wang, and X. Zhuang (2015). "Spatially resolved, highly multiplexed RNA profiling in single cells". In: *Science* 348.6233. DOI: `10.1126/science.aaa6090`. eprint: `https://science.sciencemag.org/content/348/6233/aaa6090.full.pdf` (cit. on p. 4).

Chhetri, R. K., F. Amat, Y. Wan, B. Höckendorf, W. C. Lemon, and P. J. Keller (2015). “Whole-animal functional and developmental imaging with isotropic spatial resolution”. In: *Nature Methods* 12.12, pp. 1171–1178. DOI: `10.1038/nmeth.3632` (cit. on p. 34).

Christensen, R. P. et al. (2015). “Untwisting the \textit{Caenorhabditis elegans} embryo”. In: *eLife* 4. Ed. by O. Hobert, e10070. DOI: `10.7554/eLife.10070` (cit. on p. 58).

Coles, C. H. and F. Bradke (2015). “Coordinating Neuronal Actin-Microtubule Dynamics”. In: *Current Biology* 25.15, R677–R691. DOI: `10.1016/j.cub.2015.06.020` (cit. on p. 51).

Collet, Y. (2011). *LZ4 Compression algorithm*. URL: `https://github.com/lz4/lz4/releases` (visited on 05/14/2019) (cit. on p. 24).

Cristofori-Armstrong, B. et al. (2015). “Xenopus borealis as an alternative source of oocytes for biophysical and pharmacological studies of neuronal ion channels”. In: *Scientific Reports* 5, p. 14763 (cit. on p. 81).

Daetwyler, S., U. Günther, C. D. Modes, K. Harrington, and J. Huisken (2019). “Multi-sample SPIM image acquisition, processing and analysis of vascular growth in zebrafish”. In: *Development* 146.6. DOI: `10.1242/dev.173757` (cit. on pp. 29, 55, 65, 67).

Dodt, H.-U. et al. (2007). “Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain”. In: *Nature Methods* 4, p. 331 (cit. on p. 11).

Drost, J. and H. Clevers (2018). “Organoids in cancer research”. In: *Nature Reviews Cancer* 18.7, pp. 407–418. DOI: `10.1038/s41568-018-0007-6` (cit. on p. 74).

Ellett, F., L. Pase, J. W. Hayman, A. Andrianopoulos, and G. J. Lieschke (2011). “mpeg1 promoter transgenes direct macrophage-lineage expression in zebrafish.” eng. In: *Blood* 117.4, e49–56. DOI: `10.1182/blood-2010-10-314120` (cit. on p. 101).

Erturk, A. and F. Bradke (2013). “High-resolution imaging of entire organs by 3-dimensional imaging of solvent cleared organs (3DISCO).” eng. In: *Experimental neurology* 242, pp. 57–64. DOI: `10.1016/j.expneurol.2012.10.018` (cit. on p. 4).

Felton, T., P. F. Troke, and W. W. Hope (2014). “Tissue penetration of antifungal agents.” eng. In: *Clinical microbiology reviews* 27.1, pp. 68–88. DOI: `10.1128/CMR.00046-13` (cit. on p. 86).

Flynn, K. C. (2013). “The cytoskeleton and neurite initiation”. eng. In: *Bioarchitecture* 3.4, pp. 86–109. DOI: `10.4161/bioa.26259` (cit. on p. 51).

Fuchs, E., J. S. Jaffe, R. A. Long, and F. Azam (2002). “Thin laser light sheet microscope for microbial oceanography”. In: *Opt. Express* 10.2, pp. 145–154. DOI: `10.1364/OE.10.000145` (cit. on p. 8).

Gao, R. et al. (2019). “Cortical column and whole-brain imaging with molecular contrast and nanoscale resolution”. In: *Science* 363.6424. DOI: `10.1126/science.aau8302`. eprint: `https://science.sciencemag.org/content/363/6424/eaau8302.full.pdf` (cit. on p. 4).

Google (2011). *Snappy compression algorithm*. URL: `https://github.com/google/snappy/releases` (visited on 05/14/2019) (cit. on p. 24).

Hanson, A. J. and H. Ma (1995). *Parallel Transport Approach to Curve Framing*. Tech. rep. (cit. on p. 59).

Heaster, T. M., A. J. Walsh, and M. C. Skala (2016). “Differentiating quiescent cancer cell populations in heterogeneous samples with fluorescence lifetime imaging”. In: *Proc.SPIE*. Vol. 9719 (cit. on p. 74).

Heemskerk, I. and S. J. Streichan (2015). “Tissue cartography: compressing bio-image data by dimensional reduction”. In: *Nature Methods* 12, p. 1139 (cit. on p. 61).

Hell, S. W. and J. Wichmann (1994). “Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy”. In: *Opt. Lett.* 19.11, pp. 780–782. DOI: `10.1364/OL.19.000780` (cit. on p. 4).

Hetrick, B., M. S. Han, L. A. Helgeson, and B. J. Nolen (2013). “Small molecules CK-666 and CK-869 inhibit actin-related protein 2/3 complex by blocking an activating conformational change”. eng. In: *Chemistry & biology* 20.5, pp. 701–712. DOI: `10.1016/j.chembiol.2013.03.019` (cit. on p. 63).

Hörl, D. et al. (2018). “BigStitcher: Reconstructing high-resolution image datasets of cleared and expanded samples”. In: *bioRxiv*. DOI: `10.1101/343954`. eprint: `https://www.biorxiv.org/content/early/2018/06/10/343954.full.pdf` (cit. on p. 6).

Huaman, R., C. Riviere, C. Frindel, A. Bassi, and D. Rousseau (2018). “On the use of image quality measures of multi-views in light sheet fluorescence 3D microscopy”. In: *Proc.SPIE*. Vol. 10677 (cit. on p. 36).

Huffman, D. A. (1952). “A Method for the Construction of Minimum-Redundancy Codes”. In: *Proceedings of the IRE* 40.9, pp. 1098–1101. DOI: `10.1109/JRPROC.1952.273898` (cit. on p. 24).

Huisken, J. and D. Y. R. Stainier (2007). “Even fluorescence excitation by multidirectional selective plane illumination microscopy (mSPIM)”. In: *Optics Letters* 32.17, p. 2608. DOI: `10.1364/OL.32.002608` (cit. on p. 34).

– (2009). “Selective plane illumination microscopy techniques in developmental biology”. In: *Development* 136.12, pp. 1963–1975. DOI: `10.1242/dev.022426` (cit. on p. 9).

Huisken, J., J. Swoger, F. Del Bene, J. Wittbrodt, and E. H. K. Stelzer (2004). “Optical Sectioning Deep Inside Live Embryos by Selective Plane Illumination Microscopy”. In: *Science* 305.5686, 1007 LP –1009 (cit. on p. 6).

Jesacher, A. and M. J. Booth (2011). “Sensorless adaptive optics for microscopy”. In: *Proc.SPIE*. Vol. 7931 (cit. on p. 93).

Jin, S.-W., D. Beis, T. Mitchell, J.-N. Chen, and D. Y. R. Stainier (2005). “Cellular and molecular analyses of vascular tube and lumen formation in zebrafish.” eng. In: *Development (Cambridge, England)* 132.23, pp. 5199–5209. DOI: `10.1242/dev.02087` (cit. on p. 101).

Kanungo, J., Y.-L. Zheng, B. Mishra, and H. C. Pant (2008). “Zebrafish Rohon-Beard Neuron Development: Cdk5 in the Midst”. In: *Neurochemical Research* 34.6, p. 1129. DOI: `10.1007/s11064-008-9885-4` (cit. on p. 53).

Kaufmann, A., M. Mickoleit, M. Weber, and J. Huisken (2012). “Multilayer mounting enables long-term imaging of zebrafish development in a light sheet microscope”. In: *Development* 139.17, pp. 3242–3247. DOI: `10.1242/dev.082586` (cit. on pp. 56, 57, 102).

Keller, P. J., A. D. Schmidt, J. Wittbrodt, and E. H. K. Stelzer (2008). “Reconstruction of Zebrafish Early Embryonic Development by Scanned Light Sheet Microscopy”. In: *Science* 322.5904, pp. 1065–1069. DOI: `10.1126/science.1162493` (cit. on p. 9).

Ketschek, A., S. Jones, M. Spillane, F. Korobova, T. Svitkina, and G. Gallo (2015). “Nerve growth factor promotes reorganization of the axonal microtubule array at sites of axon collateral branching.” eng. In: *Developmental neurobiology* 75.12, pp. 1441–1461. DOI: `10.1002/dneu.22294` (cit. on p. 63).

Knopf, F. et al. (2011). “Bone regenerates via dedifferentiation of osteoblasts in the zebrafish fin.” eng. In: *Developmental cell* 20.5, pp. 713–724. DOI: `10.1016/j.devcel.2011.04.014` (cit. on p. 101).

Kreyszig, E. and A. Pendl (1975). *Spherical curves and their analogues in affine differential geometry*. Vol. 48, p. 423. DOI: `10.1090/S0002-9939-1975-0365369-0` (cit. on p. 59).

Krizhevsky, A., I. Sutskever, and G. E. Hinton (2012). “ImageNet Classification with Deep Convolutional Neural Networks”. In: *Advances in Neural Information Processing Systems 25*. Ed. by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger. Curran Associates, Inc., pp. 1097–1105 (cit. on p. 95).

Krull, A., T. Buchholz, and F. Jug (2018). "Noise2Void - Learning Denoising from Single Noisy Images". In: *CoRR* abs/1811.10980. arXiv: `1811.10980` (cit. on p. 96).

Krzic, U., S. Gunther, T. E. Saunders, S. J. Streichan, and L. Hufnagel (2012). "Multiview light-sheet microscope for rapid in toto imaging". In: *Nature Methods* 9.7, pp. 730–733. DOI: `10.1038/nmeth.2064` (cit. on p. 34).

Lee, T. J., J. W. Lee, E. M. Haynes, K. W. Eliceiri, and M. C. Halloran (2017). "The Kinesin Adaptor Calsyntenin-1 Organizes Microtubule Polarity and Regulates Dynamics during Sensory Axon Arbor Development". In: *Frontiers in Cellular Neuroscience* 11, p. 107. DOI: `10.3389/fncel.2017.00107` (cit. on pp. 52, 53).

Lemire, D. and L. Boytsov (2012). "Decoding billions of integers per second through vectorization". In: *CoRR* abs/1209.2137. arXiv: `1209.2137` (cit. on p. 27).

Liu, Y. (2011). "The DIADEM and Beyond". In: *Neuroinformatics* 9.2, pp. 99–102. DOI: `10.1007/s12021-011-9102-5` (cit. on p. 57).

Long, B. et al. (2017). "SmartScope2: Simultaneous Imaging and Reconstruction of Neuronal Morphology". In: *Scientific Reports* 7.1, p. 9325. DOI: `10.1038/s41598-017-10067-w` (cit. on p. 15).

McDole, K. et al. (2018). "In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level". In: *Cell*, pp. 1–18. DOI: `10.1016/j.cell.2018.09.031` (cit. on p. 12).

Megason, S. G. (2009). "In toto imaging of embryogenesis with confocal time-lapse microscopy." eng. In: *Methods in molecular biology (Clifton, N.J.)* 546, pp. 317–332. DOI: `10.1007/978-1-60327-977-2_19` (cit. on pp. 54, 66).

Menon, S. and S. L. Gupton (2016). "Building Blocks of Functioning Brain: Cytoskeletal Dynamics in Neuronal Development." eng. In: *International review of cell and molecular biology* 322, pp. 183–245. DOI: `10.1016/bs.ircmb.2015.10.002` (cit. on p. 51).

"Method of the Year" (2014). In: *Nature Methods* 12, p. 1 (cit. on p. 11).

Miller, A. L. and W. M. Bement (2008). "Regulation of cytokinesis by Rho GTPase flux". In: *Nature Cell Biology* 11, p. 71 (cit. on p. 80).

Mizoguchi, T., H. Verkade, J. K. Heath, A. Kuroiwa, and Y. Kikuchi (2008). "Sdf1/Cxcr4 signaling controls the dorsal migration of endodermal cells during zebrafish gastrulation." eng. In: *Development (Cambridge, England)* 135.15, pp. 2521–2529. DOI: `10.1242/dev.020107` (cit. on p. 101).

Nirschl, J. J., A. E. Ghiretti, and E. L. F. Holzbaur (2017). "The impact of cytoskeletal organization on the local regulation of neuronal transport." eng. In: *Nature reviews. Neuroscience* 18.10, pp. 585–597. DOI: `10.1038/nrn.2017.100` (cit. on p. 51).

Nüsslein-Volhard, C. and R. Dahm (2001). *Zebrafish: A Practial Approach*. Vol. xviii (cit. on p. 101).

Oberhumer, M. F. (1996). *LZO Compression algorithm*. URL: `http://www.oberhumer.com/opensource/lzo/` (visited on 05/14/2019) (cit. on p. 24).

Oliveira, S. de, E. E. Rosowski, and A. Huttenlocher (2016). "Neutrophil migration in infection and wound repair: going forward in reverse." eng. In: *Nature reviews. Immunology* 16.6, pp. 378–391. DOI: `10.1038/nri.2016.49` (cit. on p. 87).

Ounkomol, C., S. Seshamani, M. M. Maleckar, F. Collman, and G. R. Johnson (2018). "Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy". In: *Nature Methods* 15.11, pp. 917–920. DOI: `10.1038/s41592-018-0111-2` (cit. on p. 95).

Pacheco, A. and G. Gallo (2016). "Actin filament-microtubule interactions in axon initiation and branching." eng. In: *Brain research bulletin* 126.Pt 3, pp. 300–310. DOI: `10.1016/j.brainresbull.2016.07.013` (cit. on p. 51).

Park, J.-H., L. Kong, Y. Zhou, and M. Cui (2017). "Large-field-of-view imaging by multi-pupil adaptive optics". In: *Nature Methods* 14, p. 581 (cit. on p. 93).

Peng, H., F. Long, T. Zhao, and E. Myers (2011). "Proof-editing is the Bottleneck Of 3D Neuron Reconstruction: The Problem and Solutions". In: *Neuroinformatics* 9.2, pp. 103–105. DOI: `10.1007/s12021-010-9090-x` (cit. on p. 57).

Pietzsch, T., S. Saalfeld, S. Preibisch, and P. Tomancak (2015). "BigDataViewer: visualization and processing for large image data sets". In: *Nature Methods* 12, p. 481 (cit. on p. 11).

Ponomareva, O. Y., I. C. Holmen, A. J. Sperry, K. W. Eliceiri, and M. C. Halloran (2014). "Calsyntenin-1 regulates axon branching and endosomal trafficking during sensory neuron development in vivo." eng. In: *The Journal of neuroscience : the official journal of the Society for Neuroscience* 34.28, pp. 9235–9248. DOI: `10.1523/JNEUROSCI.0561-14.2014` (cit. on p. 52).

Preibisch, S., F. Amat, E. Stamataki, M. Sarov, R. H. Singer, E. Myers, and P. Tomancak (2014). "Efficient Bayesian-based multiview deconvolution". In: *Nature Methods* 11, p. 645 (cit. on p. 48).

Preibisch, S., S. Saalfeld, J. Schindelin, and P. Tomancak (2010). "Software for bead-based registration of selective plane illumination microscopy data". In: *Nature Methods* 7.6, pp. 418–419. DOI: `10.1038/nmeth0610-418` (cit. on pp. 44, 102).

Püspöki, Z., M. Storath, D. Sage, and M. Unser (2016). "Transforms and Operators for Directional Bioimage Analysis: A Survey". In: *Focus on Bio-Image Informatics*. Ed.

by W. De Vos, S. Munck, and J.-P. Timmermans. Vol. 219. Advances in Anatomy, Embryology and Cell Biology. Springer International Publishing. Chap. 3, pp. 69–93 (cit. on p. 62).

Reyes, R., M. Haendel, D. Grant, E. Melancon, and J. S. Eisen (2004). "Slow degeneration of zebrafish Rohon-Beard neurons during programmed cell death." eng. In: *Developmental dynamics : an official publication of the American Association of Anatomists* 229.1, pp. 30–41. DOI: `10.1002/dvdy.10488` (cit. on p. 53).

Robinson, A. H. and C. Cherry (1967). "Results of a prototype television bandwidth compression scheme". In: *Proceedings of the IEEE* 55.3, pp. 356–364. DOI: `10.1109/PROC.1967.5493` (cit. on p. 24).

Ronneberger, O., P. Fischer, and T. Brox (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *CoRR* abs/1505.04597. arXiv: `1505.04597` (cit. on p. 94).

Rosowski, E. E., B. P. Knox, L. S. Archambault, A. Huttenlocher, N. P. Keller, R. T. Wheeler, and J. M. Davis (2018). "The Zebrafish as a Model Host for Invasive Fungal Infections." eng. In: *Journal of fungi (Basel, Switzerland)* 4.4. DOI: `10.3390/jof4040136` (cit. on p. 87).

Royer, L. A., W. C. Lemon, R. K. Chhetri, Y. Wan, M. Coleman, E. W. Myers, and P. J. Keller (2016). "Adaptive light-sheet microscopy for long-term, high-resolution imaging in living organisms". In: *Nature Biotechnology* 34.12, pp. 1267–1278. DOI: `10.1038/nbt.3708` (cit. on pp. 14, 36, 38, 42).

Saalfeld, S., A. Cardona, V. Hartenstein, and P. Tomančák (2010). "As-rigid-as-possible mosaicking and serial section registration of large ssTEM datasets". In: *Bioinformatics* 26.12, pp. 57–63. DOI: `10.1093/bioinformatics/btq219` (cit. on p. 42).

Scherf, N. and J. Huisken (2015). "The smart and gentle microscope". In: *Nature Biotechnology* 33.8, pp. 815–818. DOI: `10.1038/nbt.3310` (cit. on pp. 13, 96).

Schindelin, J. et al. (2012). "Fiji: an open-source platform for biological-image analysis". In: *Nature Methods* 9, p. 676 (cit. on p. 6).

Schmid, B. and J. Huisken (2015). "Real-time multi-view deconvolution". In: *Bioinformatics* 31.20, pp. 3398–3400. DOI: `10.1093/bioinformatics/btv387` (cit. on pp. 48, 70).

Schmid, B., P. Tripal, T. Fraaß, C. Kersten, B. Ruder, A. Grüneboom, J. Huisken, and R. Palmisano (2019). "3Dscript: animating 3D/4D microscopy data using a natural-language-based syntax". In: *Nature Methods* 16.4, pp. 278–280. DOI: `10.1038/s41592-019-0359-1` (cit. on p. 11).

Schmid, B. et al. (2013). “High-speed panoramic light-sheet microscopy reveals global endodermal cell dynamics”. In: *Nature Communications* 4, pp. 1–10. doi: `10.1038/ncomms3207` (cit. on pp. 14, 20, 21, 34, 35, 61, 83).

Schmied, C., P. Steinbach, T. Pietzsch, S. Preibisch, and P. Tomancak (2015). “An automated workflow for parallel processing of large multiview SPIM recordings”. In: *Bioinformatics* 32.7, pp. 1112–1114. doi: `10.1093/bioinformatics/btv706` (cit. on pp. 12, 29, 55).

Seward, J. (1996). *Bzip2 compression algorithm*. url: `http://www.bzip.org/` (visited on 05/14/2019) (cit. on p. 24).

Shah, G. (2016). “Pan-embryo Cell Dynamics”. PhD thesis. Technical University of Dresden (cit. on pp. 65, 67).

Shannon, C. E. (1948). “A Mathematical Theory of Communication”. In: *Bell System Technical Journal* 27.3, pp. 379–423. doi: `10.1002/j.1538-7305.1948.tb01338.x`. arXiv: `9411012 [chao-dyn]` (cit. on pp. 38, 42).

Sharick, J. T., J. J. Jeffery, M. R. Karim, C. M. Walsh, K. Esbona, R. S. Cook, and M. C. Skala (2019). “Cellular Metabolic Heterogeneity In Vivo Is Recapitulated in Tumor Organoids”. In: *Neoplasia* 21.6, pp. 615–626. doi: `https://doi.org/10.1016/j.neo.2019.04.004` (cit. on p. 74).

Siedentopf, H. and R. Zsigmondy (1902). “Uber Sichtbarmachung und Größenbestimmung ultramikoskopischer Teilchen, mit besonderer Anwendung auf Goldrubingläser”. In: *Annalen der Physik* 315.1, pp. 1–39. doi: `10.1002/andp.19023150102` (cit. on p. 8).

Skala, M. C., K. M. Riching, A. Gendron-Fitzpatrick, J. Eickhoff, K. W. Eliceiri, J. G. White, and N. Ramanujam (2007). “In vivo multiphoton microscopy of NADH and FAD redox states, fluorescence lifetimes, and cellular morphology in precancerous epithelia”. In: *Proceedings of the National Academy of Sciences* 104.49, pp. 19494–19499. doi: `10.1073/pnas.0708425104` (cit. on p. 74).

Swoger, J., P. Verveer, K. Greger, J. Huisken, and E. H. Stelzer (2007). “Multi-view image fusion improves resolution in three-dimensional microscopy”. In: *Optics Express* 15.13, p. 8029. doi: `10.1364/OE.15.008029` (cit. on p. 34).

Tinevez, J.-Y. (2017). *Directionality*. url: `https://imagej.net/Directionality` (visited on 05/29/2019) (cit. on p. 62).

Tomer, R., K. Khairy, F. Amat, and P. J. Keller (2012). “Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy”. In: *Nature Methods* 9, p. 755 (cit. on p. 34).

Voie, A. H., D. H. Burns, and F. A. Spelman (1993). "Orthogonal-plane fluorescence optical sectioning: Three-dimensional imaging of macroscopic biological specimens". In: *Journal of Microscopy* 170.3, pp. 229–236. DOI: `10.1111/j.1365-2818.1993.tb03346.x` (cit. on p. 8).

Walsh, A. J., R. S. Cook, H. C. Manning, D. J. Hicks, A. Lafontant, C. L. Arteaga, and M. C. Skala (2013). "Optical metabolic imaging identifies glycolytic levels, subtypes, and early-treatment response in breast cancer." eng. In: *Cancer research* 73.20, pp. 6164–6174. DOI: `10.1158/0008-5472.CAN-13-0527` (cit. on p. 74).

Weber, M., M. Mickoleit, and J. Huisken (2014). "Multilayer Mounting for Long-term Light Sheet Microscopy of Zebrafish". In: *JoVE* 84, e51119. DOI: `doi:10.3791/51119` (cit. on p. 57).

Weber, M., N. Scherf, A. M. Meyer, D. Panáková, P. Kohl, and J. Huisken (2017). "Cell-accurate optical mapping across the entire developing heart". In: *eLife* 6. Ed. by D. Yelon, e28307. DOI: `10.7554/eLife.28307` (cit. on p. 58).

Weigert, M. et al. (2018). "Content-aware image restoration: pushing the limits of fluorescence microscopy". In: *Nature Methods* 15.12, pp. 1090–1097. DOI: `10.1038/s41592-018-0216-7` (cit. on p. 95).

Wolff, C. et al. (2018). "Multi-view light-sheet imaging and tracking with the MaMuT software reveals the cell lineage of a direct developing arthropod limb". In: *eLife* 7. Ed. by A. Sánchez Alvarado, e34410. DOI: `10.7554/eLife.34410` (cit. on p. 96).

Wu, Y. et al. (2011). "Inverted selective plane illumination microscopy (iSPIM) enables coupled cell identity lineaging and neurodevelopmental imaging in Caenorhabditis elegans". In: *Proceedings of the National Academy of Sciences* 108.43, pp. 17708–17713. DOI: `10.1073/pnas.1108494108` (cit. on p. 66).

Yip, Y. Y., S. Pernigo, A. Sanger, M. Xu, M. Parsons, R. A. Steiner, and M. P. Dodding (2016). "The light chains of kinesin-1 are autoinhibited". eng. In: *Proceedings of the National Academy of Sciences of the United States of America* 113.9, pp. 2418–2423. DOI: `10.1073/pnas.1520817113` (cit. on p. 52).

Yoo, S. K., Q. Deng, P. J. Cavnar, Y. I. Wu, K. M. Hahn, and A. Huttenlocher (2010). "Differential regulation of protrusion and polarity by PI3K during neutrophil motility in live zebrafish." eng. In: *Developmental cell* 18.2, pp. 226–236. DOI: `10.1016/j.devcel.2009.11.015` (cit. on p. 101).

DECLARATION

Declaration according to §5.5 of the doctorate regulations

I herewith declare that I have produced this paper without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This paper has not previously been presented in identical or similar form to any other German or foreign examination board.
The thesis work was conducted from 15/10/2015 to 15/07/2019 under the supervision of Dr. Jan Huisken at the Max Planck Institute of Molecular Cell Biology and Genetics and Morgridge Institute for Research.
I declare that I have not undertaken any previous unsuccessful doctorate proceedings.
I declare that I recognize the doctorate regulations of the Fakultät für Mathematik und Naturwissenschaften of the Technische Universität Dresden.

Erklärung entsprechend §5.5 der Promotionsordnung
Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.
Die Dissertation wurde im Zeitraum vom 15.10.2015 bis 15.07.2019 verfasst und von Dr. Jan Huisken, Max Planck Institut für Molekulare Zellbiologie und Genetik und Morgridge Institute for Research, betreut.
Meine Person betreffend erkläre ich hiermit, dass keine früheren erfolglosen Promotionsverfahren stattgefunden haben. Ich erkenne die Promotionsordnung der Fakultät für Mathematik und und Naturwissenschaften, Technische Universität Dresden an.

Dresden, July 2019

Jiaye He

Date